Molecular Dynamics Simulations Fundamental aspects & examples

Stefano Leoni

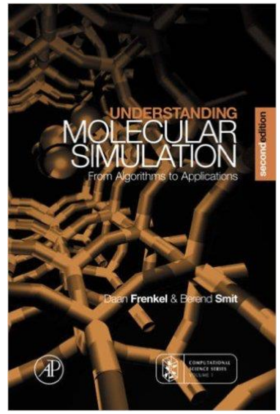
Cardiff University

leonis@cf.ac.uk

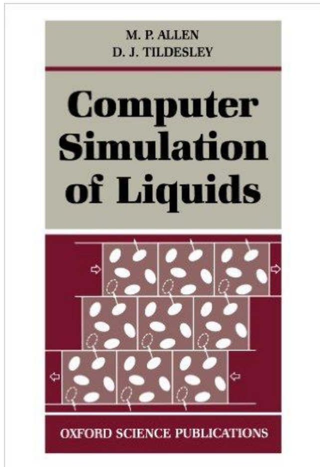
Learning Objectives

- Obtain an understanding of fundamental principles and techniques of molecular dynamics;
- Learn how to extract thermodynamic properties from molecular dynamics simulations;
- Identify what class of problems can be tackled with molecular dynamics;
- Understand current problems of molecular dynamics simulations;

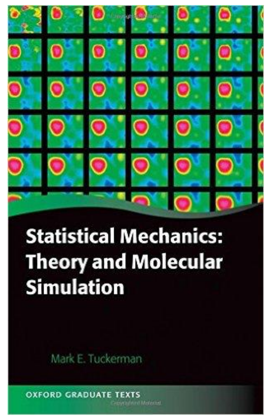
Literature



http://www.amazon.co.uk/Understanding-Molecular-Simulation-Applications-Computational/dp/0122673514



http://www.amazon.co.uk/Computer-Simulation-Liquids-Science-Publications/dp/0198556454/ref=pd_sim_14_2?ie=UTF8&dpID=51Y61tsyY3L&dpSrc=sims&preST=_AC_UL160_SR108%2C160_&refRID=134XY64QNG2KBJYY3BG1



http://www.amazon.co.uk/Statistical-Mechanics-Molecular-Simulation-Graduate/dp/0198525265

Overview

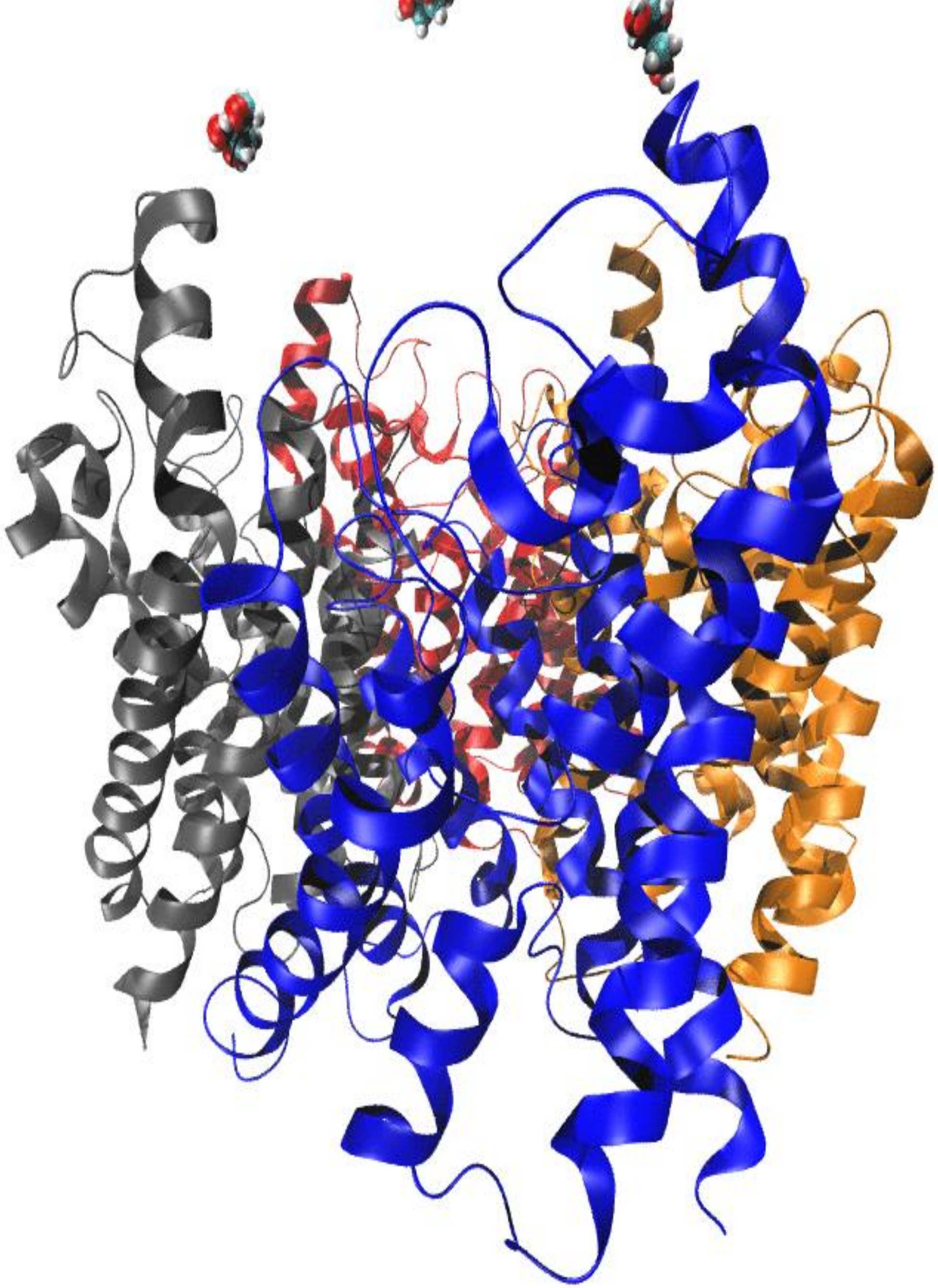
Foundations
- t-dependent Schrödinger Equation
- Derivation of a Nuclear Dynamics

Calculation of Interatomic Forces
- Potential Energy Surface
- Force Models

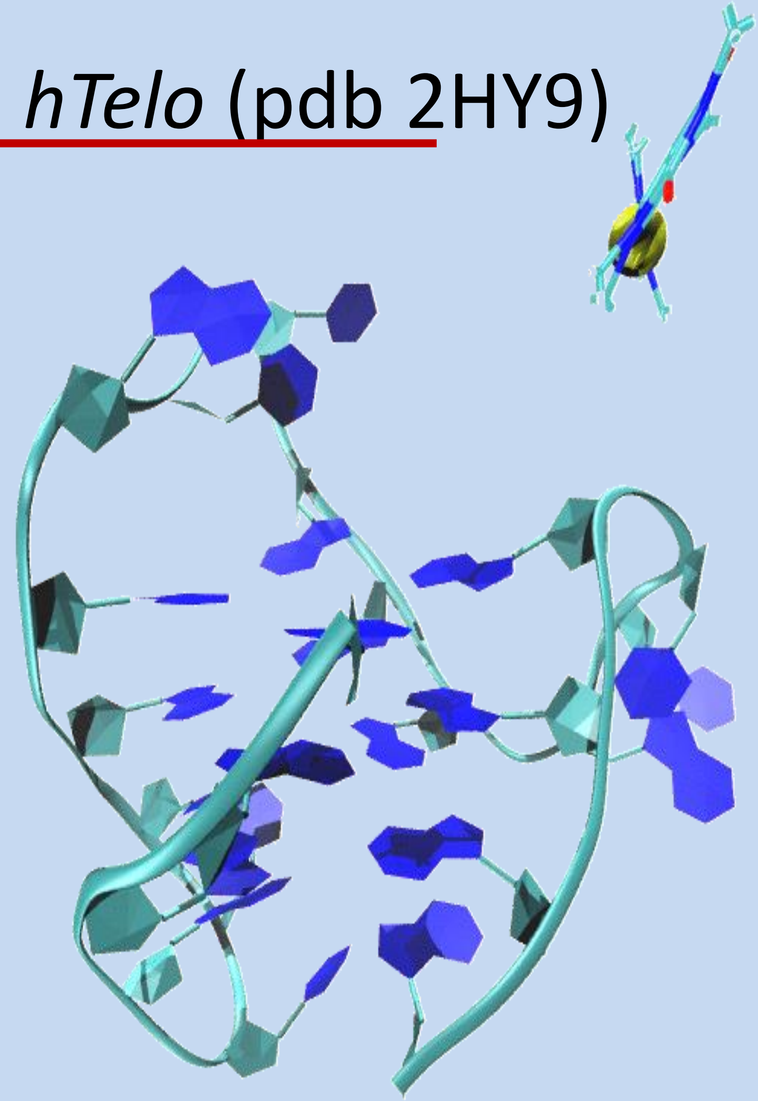
Molecular Dynamics Simulations
- \- Equation of motion & numerical solutions
- \- Evaluation of thermodynamic properties

Advanced Methods
- Metadynamics
- Transition Path Sampling

Examples
- Water
- Reconstructive Phase Transitions

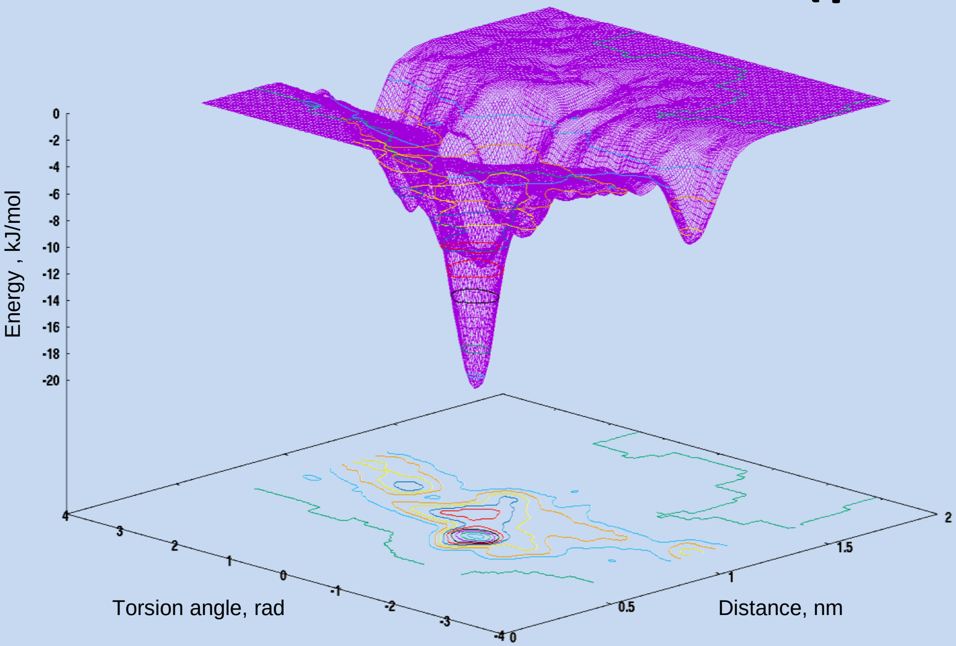


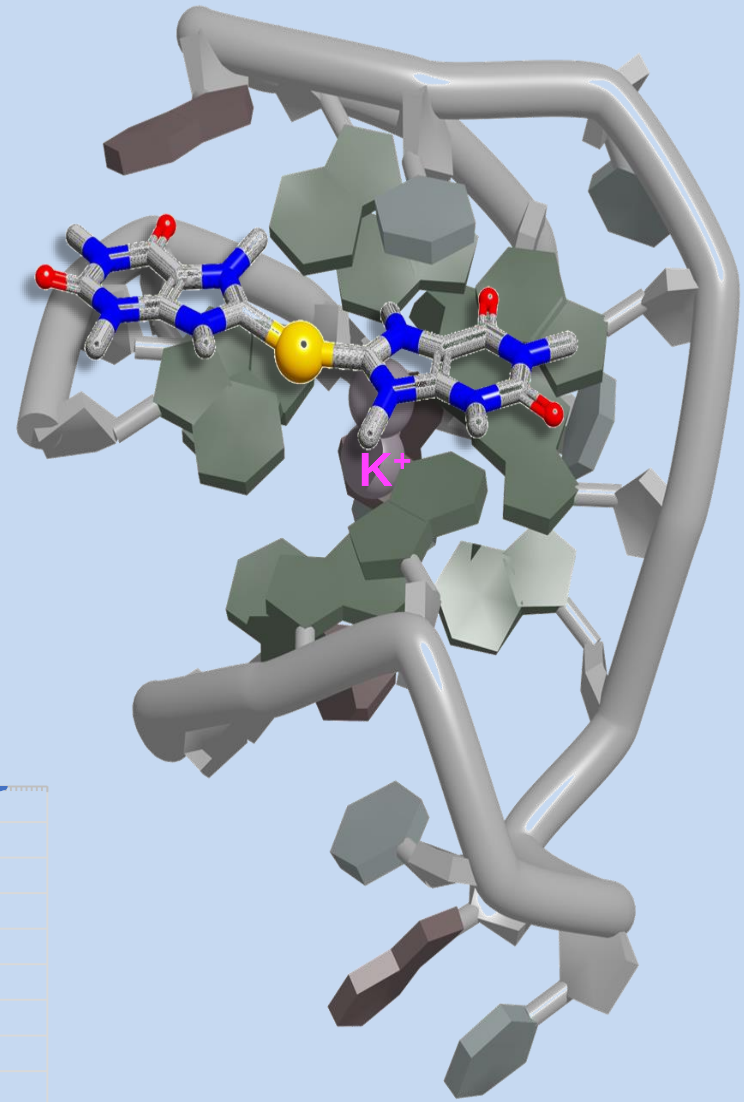
Bis-carbene with *hTelo* (pdb 2HY9)

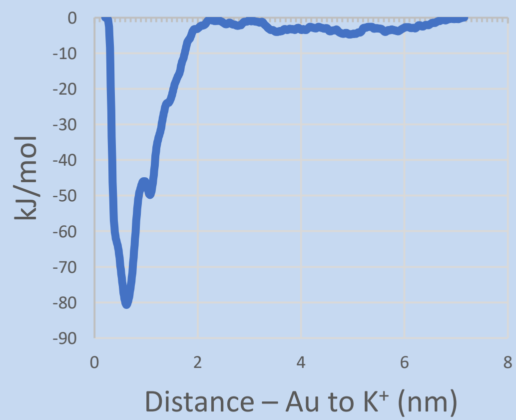


50 ns trajectory
1fs timestep
Temp - 300K
Pressure - 1atm
Solvent – H_2O (not shown)
Ion concentration – 0.15M KCl (not shown)

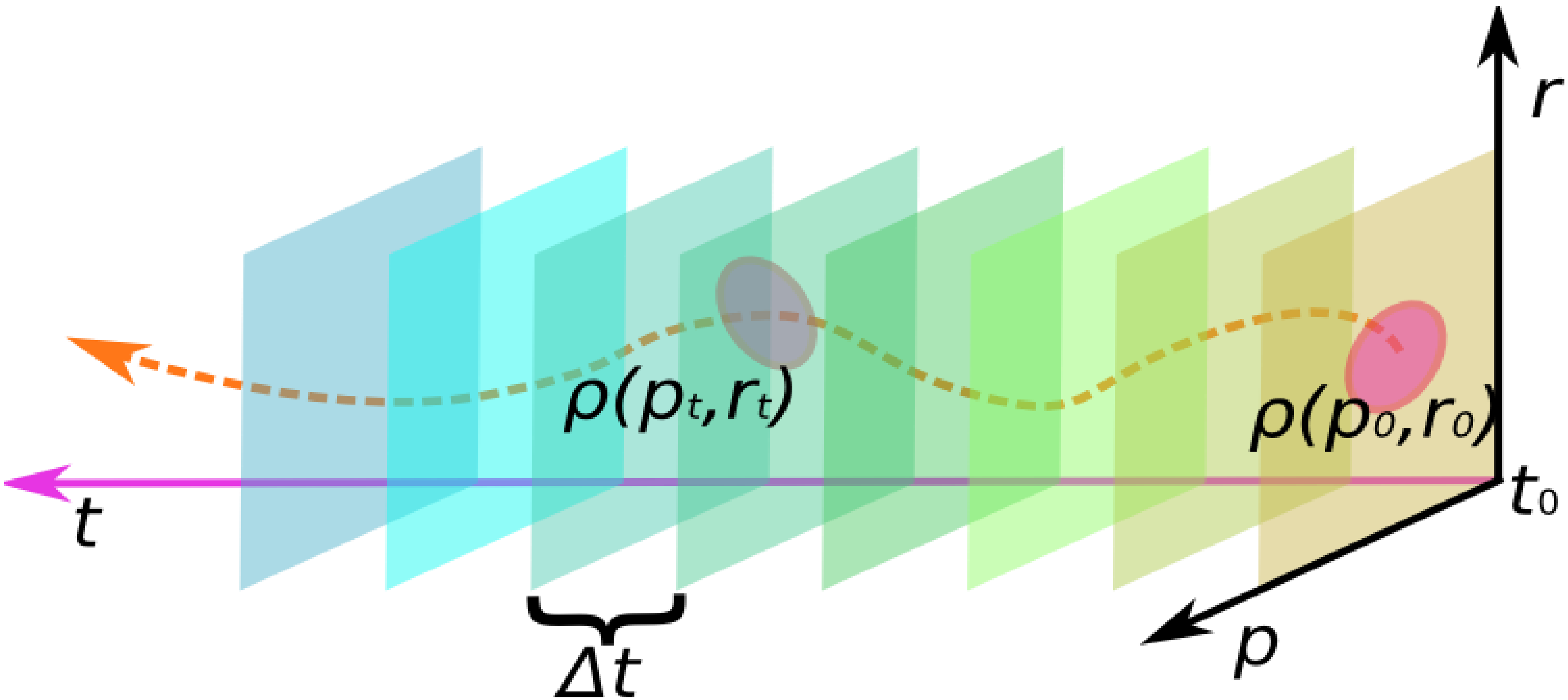
Free energy surface (FES) – bis carbene with *hTelo* (pdb 2HY9)







MD, Trajectory



Forces ← Potential

Definition:

In Molecular Dynamics, a *trajectory* is the path (positions) left beyond by a system as a function of time. It consist of a series of *snapshots*, which are instantaneous configurations of atoms.

Molecular Dynamics - MD

Numerical method (statistical mechanics).

Newton equation of motions, *numerical* integration (Ex. Verlet algorithm).

Discrete number of particles (ensembles), finite time integration step.

Per particle (atom, rigid molecule), positions and velocities are available.

Average kinetic energy per degree of freedom $$\left\langle \frac{1}{2} m v_\alpha^2 \right\rangle = \frac{1}{2} k_B T$$

Calculation of (instantaneous) temperature from kinetic energy:

$$T(t) = \sum_{i=1}^{N} \frac{m_i v_i^2(t)}{k_B N_f}$$

$$N_f = 3N - 3$$

fluctuations $\sim \frac{1}{\sqrt{N_f}}$

1

- *Forces* on atoms
- MD from *Schrödinger* Equation
- Calculation of the forces
- MD flavors (Classical, BO, CP, …)

- D. Marx, J. Hutter, Ab initio molecular dynamics: Theory and Implementation, *Modern Methods and Algorithms of Quantum Chemistry*, J. Grotendorst (Ed.), John von Neumann Institute for Computing, Jülich, NIC Series, Vol. 1, ISBN 3-00-005618-1, pp. 301-449, 2000.

MD from Schrödinger Equation

$$i\hbar \frac{\partial}{\partial t}\Phi(\{\mathbf{r}_i\},\{\mathbf{R}_I\};t) = \mathcal{H}\Phi(\{\mathbf{r}_i\},\{\mathbf{R}_I\};t)$$

t-dependent SE

$$\mathcal{H} = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_i \frac{\hbar^2}{2m_e}\nabla_i^2$$

$$+\frac{1}{4\pi\varepsilon_0}\sum_{i<j}\frac{e^2}{|r_i - r_j|} - \frac{1}{4\pi\varepsilon_0}\sum_{I,i}\frac{e^2 Z_I}{|R_I - r_i|} + \frac{1}{4\pi\varepsilon_0}\sum_{I<J}\frac{e^2 Z_I Z_J}{|R_I - R_J|}$$

$$= -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + V_{n-e}(\{r_i\},\{R_I\})$$

$$= -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \mathcal{H}_e(\{r_i\},\{R_I\})$$

Electronic Hamiltonian, depends on r&R (→ $\mathcal{H}_e(\{r_i\},\{R_I\})$)

Nuclei, kinetic energy (→ $-\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2$)

Equation of Motion (nuclei)

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \int d\mathbf{r}\, \Psi^\star \mathcal{H}_\mathrm{e} \Psi$$

$$= -\nabla_I V_\mathrm{e}^\mathrm{E}(\{\mathbf{R}_I(t)\})$$

(reads as: M_I x a_I = F_I)

Further Simplification

$$V_{\mathrm{e}}^{\mathrm{E}} = \int d\mathbf{r}\, \Psi_0^{\star} \mathcal{H}_{\mathrm{e}} \Psi_0 \equiv E_0(\{\mathbf{R}_I\})$$

„Clamped Nuclei“

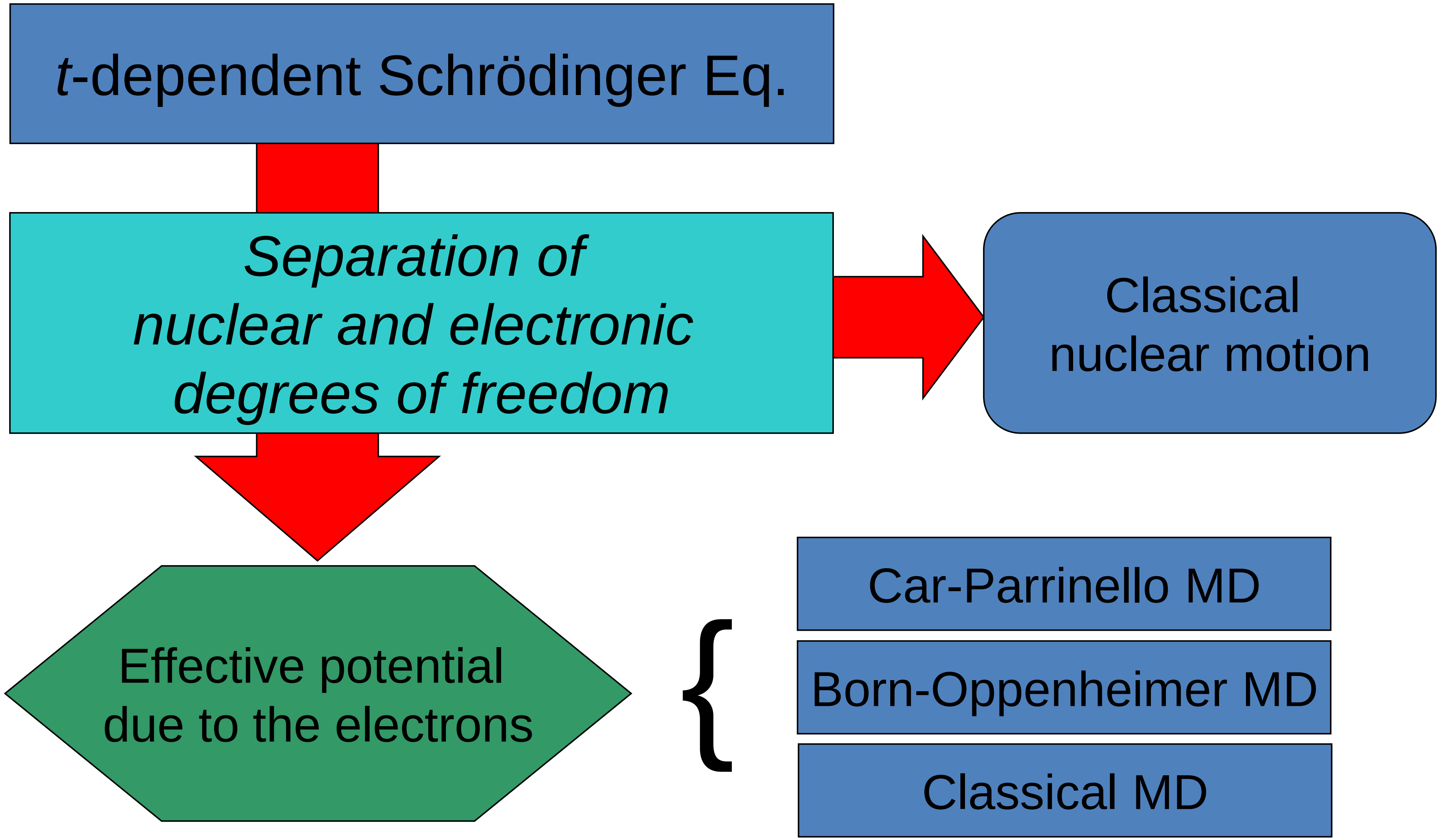
Ground state wave function

$$\mathcal{H}_{\mathrm{e}} \Psi_0 = E_0 \Psi_0$$

Ground state wave function → BO Approximation

Born Oppenheimer MD

Molecular Dynamics



Trajectory computation

Decouple Dynamics and Electronics

Compute global potential energy E_0
& derive gradients (forces)

Collect trajectories
on this potential energy surface

Approximation

- Electrons adiabatically follow nuclear motion;
- Nuclei evolve on a single BO potential energy surface (PES);
- The PES can be further approximated by „simpler“, often 2-body pair potentials.

BO Molecular Dynamics

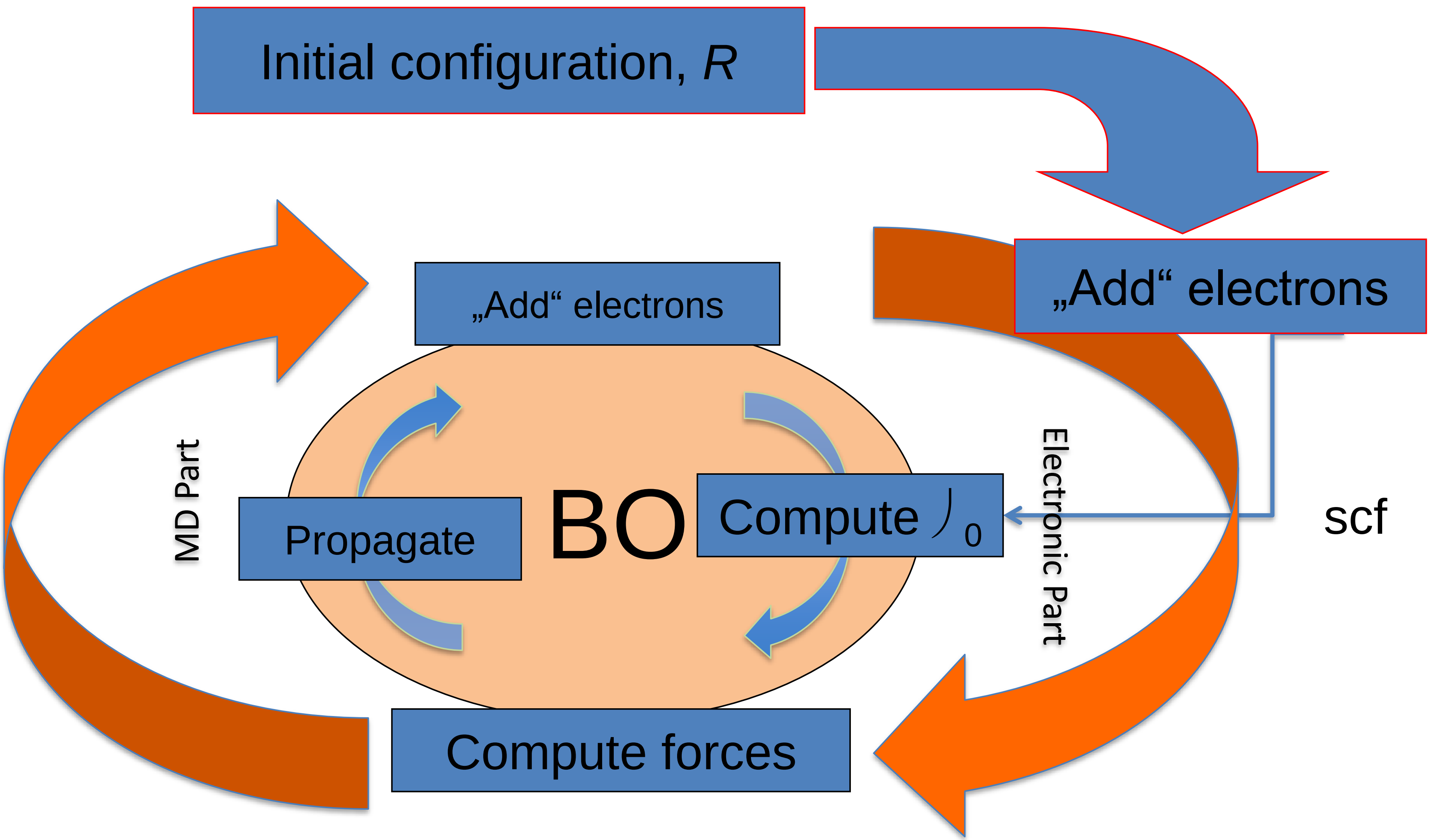
- Solve the *static (t-independent)* SE at each step;
- Given set of positions at time $t{=}t_0$;

$$M_I \ddot{R}_I(t) = -\nabla_I min_{\psi_0}\{\langle\psi_0|\mathcal{H}_e|\psi_0\rangle\}$$

$$E_o\psi_0 = \mathcal{H}_e\psi_0$$

- Minimum of $\langle\mathcal{H}_e\rangle$ has to be reached at each step.

Computational "Scheme"



Car-Parrinello MD

- Map the electronic & nuclear equation of motion into 2 classical equation of motions (classical also for the electrons)

- Use some ideas of Lagrangian mechanics

- Functional derivatives with respect to *orbitals*

Lagrangian Formulation

$$\underbrace{\sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \sum_i \frac{1}{2} \mu_i \left\langle \dot{\psi}_i \middle| \dot{\psi}_i \right\rangle}_{\text{kinetic energy}} - \underbrace{\langle \Psi_0 | \mathcal{H}_\mathrm{e} | \Psi_0 \rangle}_{\text{potential energy}} + \underbrace{constraints}_{\text{orthonormality}}$$

$$\Psi_0 = \sum_i \psi_i \qquad \sum_{i,j} \lambda_{i,j} \left(\left\langle \psi_i \middle| \psi_j \right\rangle - \delta_{ij} \right)$$

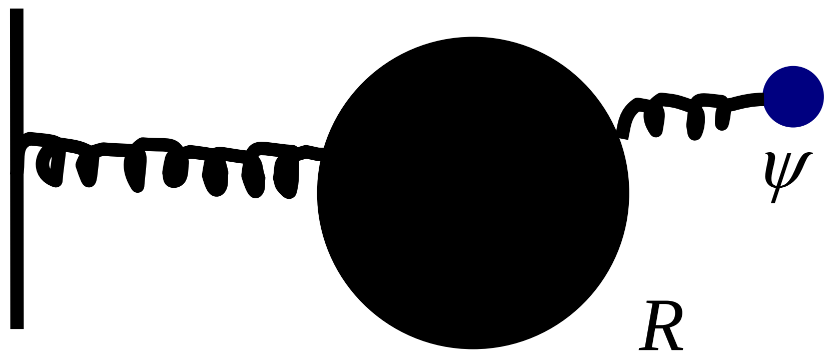
Propagate orbitals as classical objects,
with a mass and a temperature.

Equations of motions

$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I} \langle \Psi_0 | \mathcal{H}_\mathrm{e} | \Psi_0 \rangle + \frac{\partial}{\partial \mathbf{R}_I} \{constraints\}$$

$$\mu_i \ddot{\psi}_i(t) = -\frac{\delta}{\delta \psi_i^\star} \langle \Psi_0 | \mathcal{H}_\mathrm{e} | \Psi_0 \rangle + \frac{\delta}{\delta \psi_i^\star} \{constraints\}$$

Analogue



System with 2 degrees of freedom

Mass and Temperature

„Warm“ nuclei, „cold“ electrons:

$$\propto \sum_I M_I \dot{R}_I^2$$

physical temperature

$$\propto \sum_i \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle$$

„fake“ temperature

Electrons „follow“ nuclei → close to the BO pot surface

→ $\mu \approx 0$

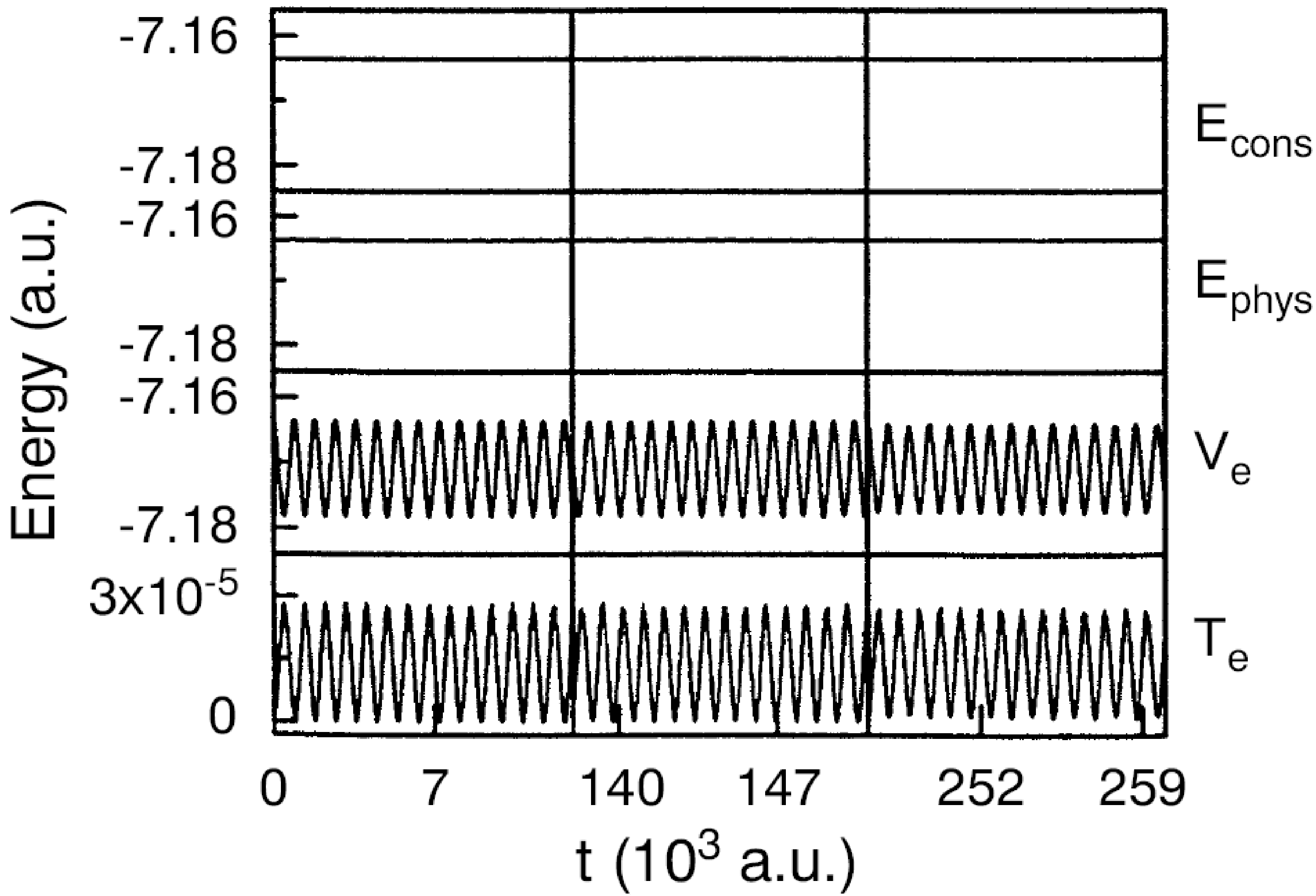
Conserved Quantitites

$$E_{\mathrm{cons}} = \sum_i \frac{1}{2}\mu_i \left\langle \dot{\psi}_i \middle| \dot{\psi}_i \right\rangle + \sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \langle \Psi_0 | \mathcal{H}_{\mathrm{e}} | \Psi_0 \rangle$$

$$E_{\mathrm{phys}} = \sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \langle \Psi_0 | \mathcal{H}_{\mathrm{e}} | \Psi_0 \rangle = E_{\mathrm{cons}} - T_{\mathrm{e}}$$

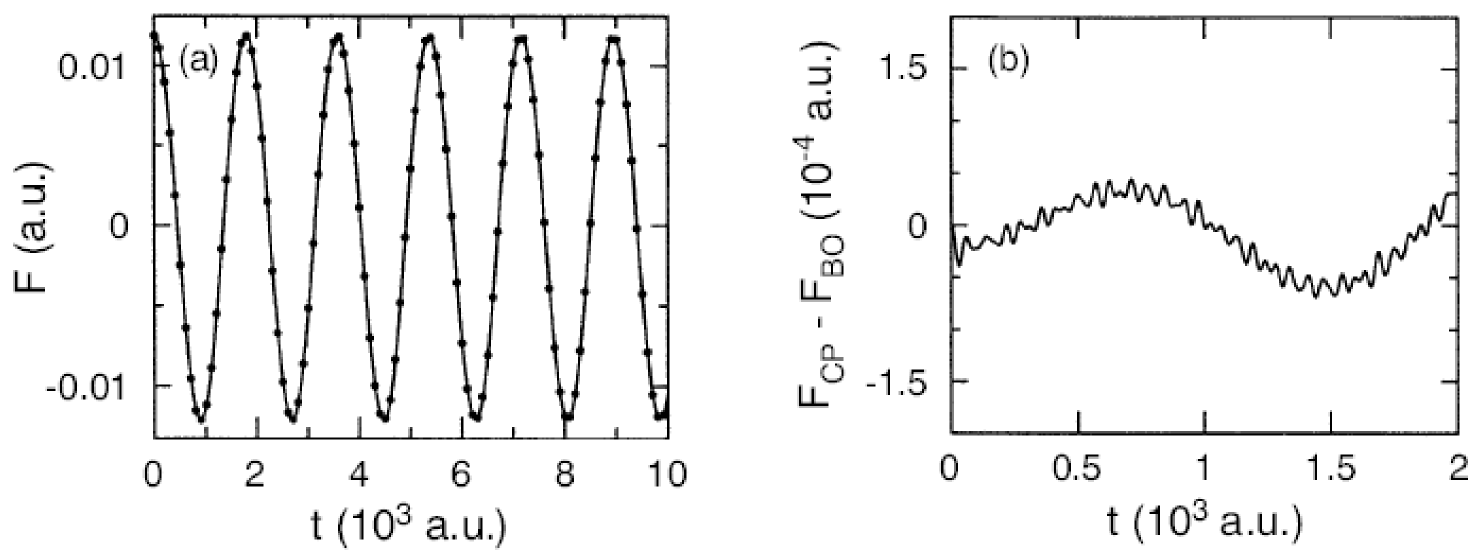
$$V_{\mathrm{e}} = \langle \Psi_0 | \mathcal{H}_{\mathrm{e}} | \Psi_0 \rangle$$

$$T_{\mathrm{e}} = \sum_i \frac{1}{2}\mu_i \left\langle \dot{\psi}_i \middle| \dot{\psi}_i \right\rangle$$



Energy (a.u.)
-7.16
-7.18
-7.16
-7.18
-7.16
-7.18
3x10-5
0
E cons
E phys
V e
T e
0
7
140
147
252
259
t (10^3 a.u.)

Comparison BO, CP

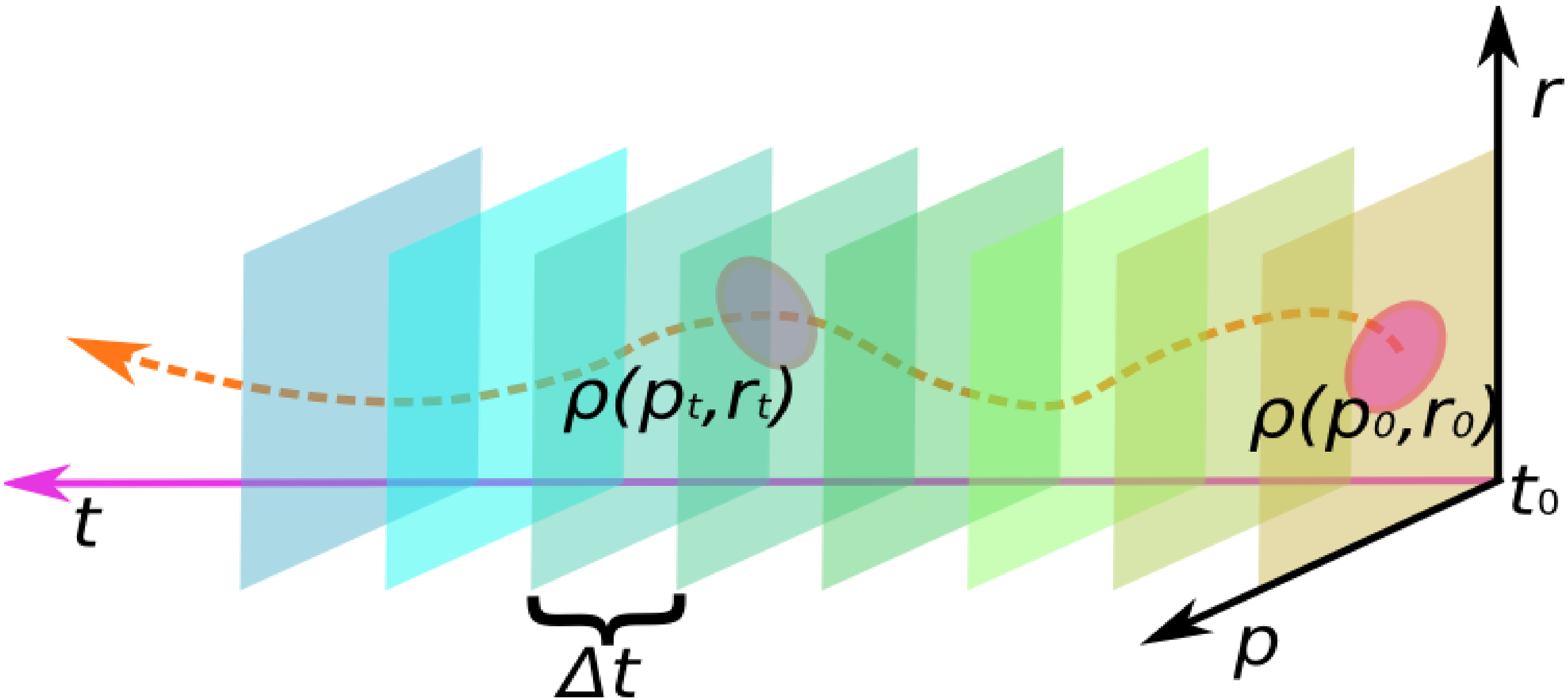


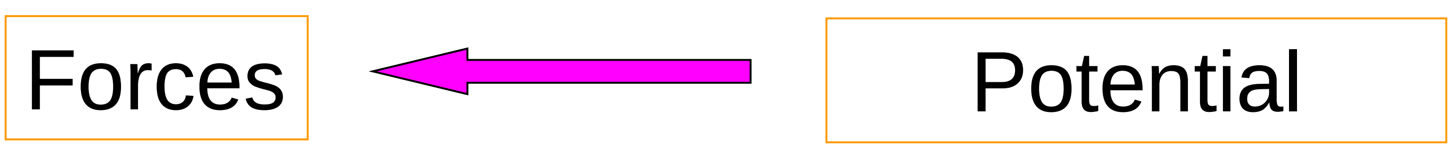
$$V_{\mathrm{e}}^{\mathrm{E}} \approx V_{\mathrm{e}}^{\mathrm{approx}}(\{\mathbf{R}_I\}) = \sum_{I=1}^{N} v_1(\mathbf{R}_I) + \sum_{I<J}^{N} v_2(\mathbf{R}_I, \mathbf{R}_J) + \sum_{I<J<K}^{N} v_3(\mathbf{R}_I, \mathbf{R}_J, \mathbf{R}_K) + \cdots$$

Classical „Limit“

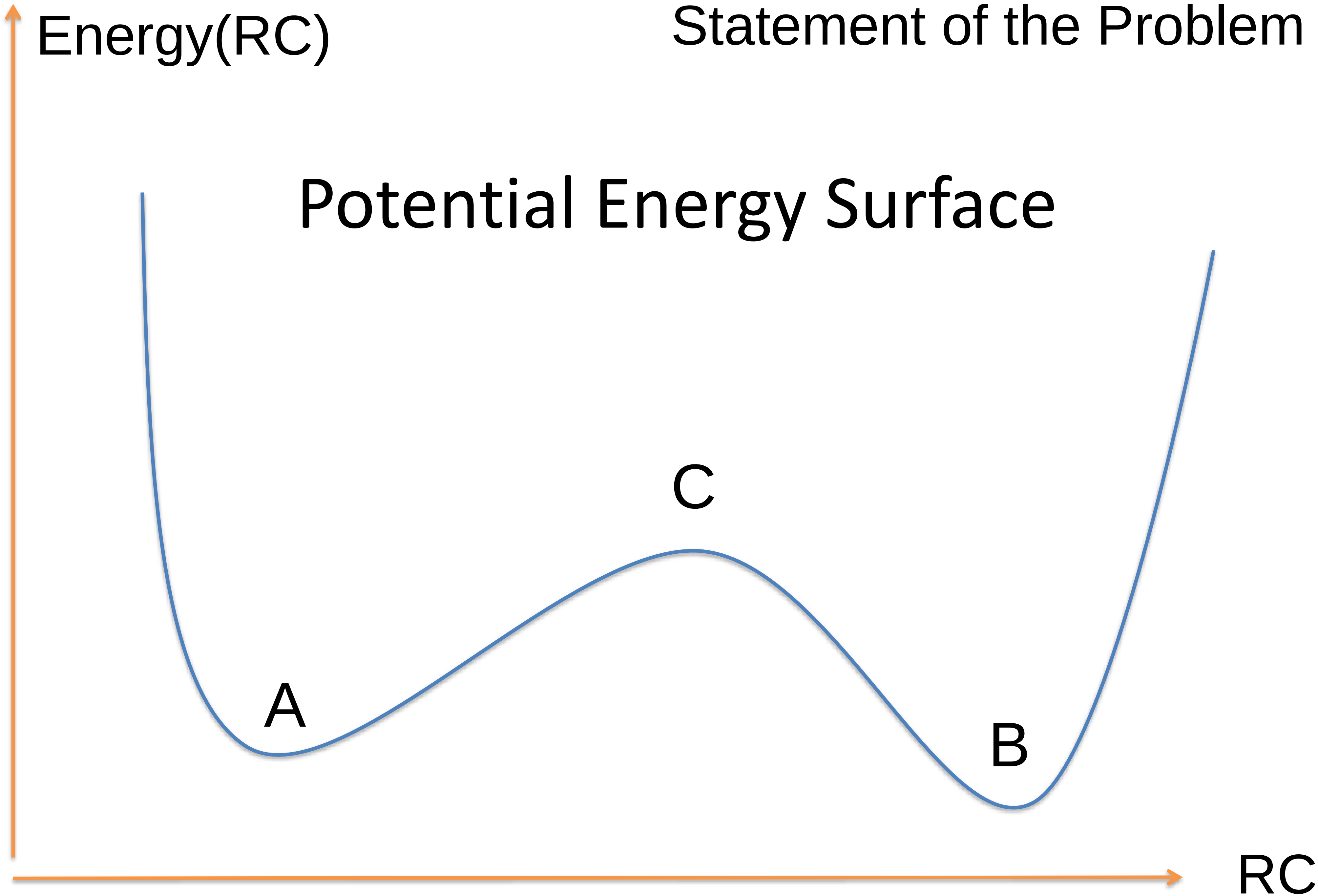
- Often pair potentials (v2) will suffice;
- Electronic degrees of freedom are not *explicitly* present;
- Fitting approaches, nowadays neural networks/ML.

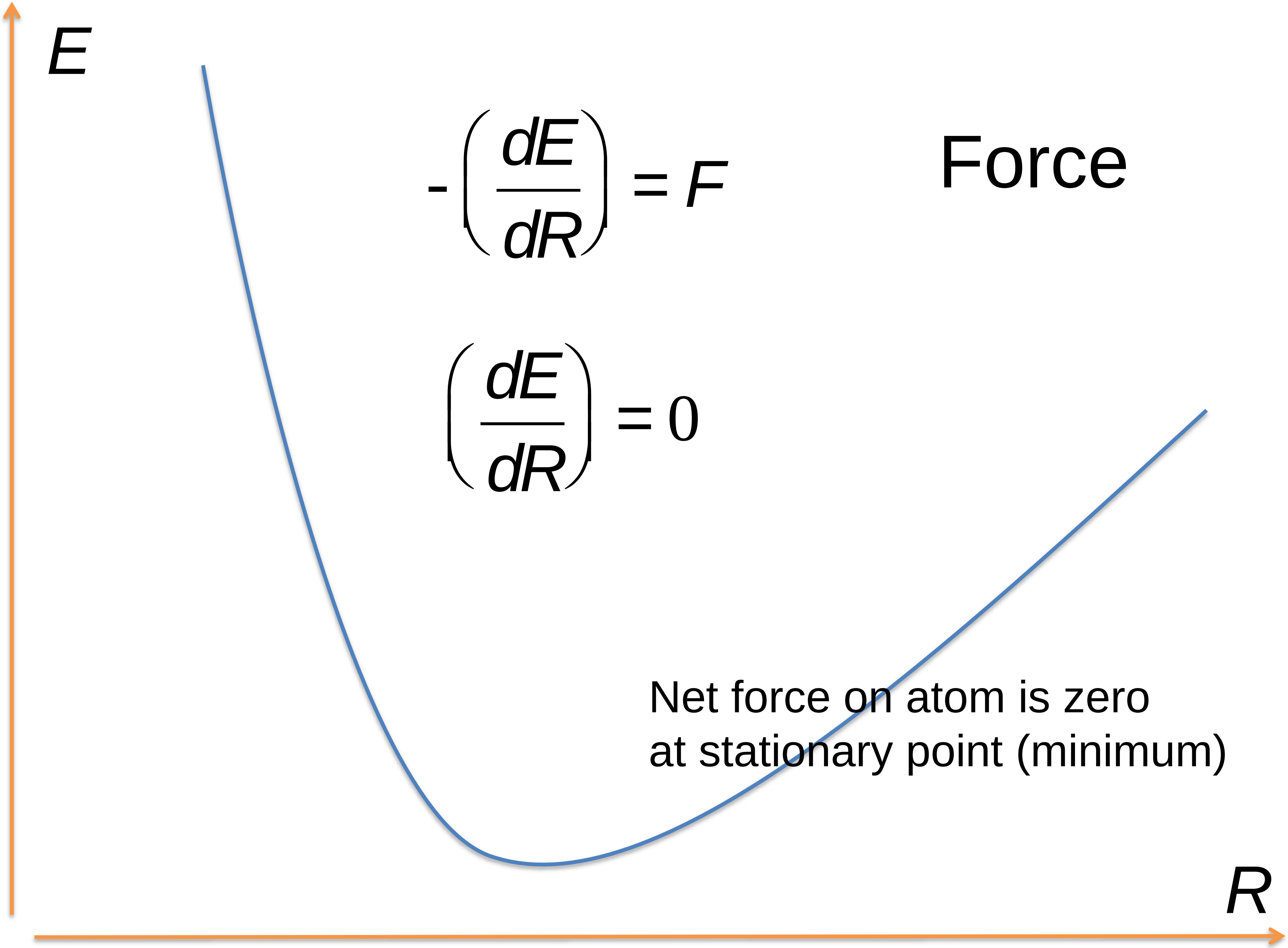
MD, Trajectory





Statement of the Problem





E
$-\left(\frac{dE}{dR}\right) = F$
Force
$\left(\frac{dE}{dR}\right) = 0$
Net force on atom is zero
at stationary point (minimum)
R

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle_{\text{Ehrenfest}}$$

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \min_{\Psi_0} \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle_{\text{BO}}$$

$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I} \langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle + \ldots$$

CP

Hellmann-Feynman Forces

$$F = -\nabla_I \langle \Psi_0 | \mathscr{H}_e | \Psi_0 \rangle$$

$$\nabla_I \langle \Psi_0 | \mathscr{H}_e | \Psi_0 \rangle = \langle \nabla_I \Psi_0 | \mathscr{H}_e | \Psi_0 \rangle + \langle \Psi_0 | \nabla_I \mathscr{H}_e | \Psi_0 \rangle + \langle \Psi_0 | \mathscr{H}_e | \nabla_I \Psi_0 \rangle$$

$$F^{HFT} = -\langle \Psi_0 | \nabla_I \mathscr{H}_e | \Psi_0 \rangle$$

For complete basis sets!

Equation of motion in *Cartesian coordinates*

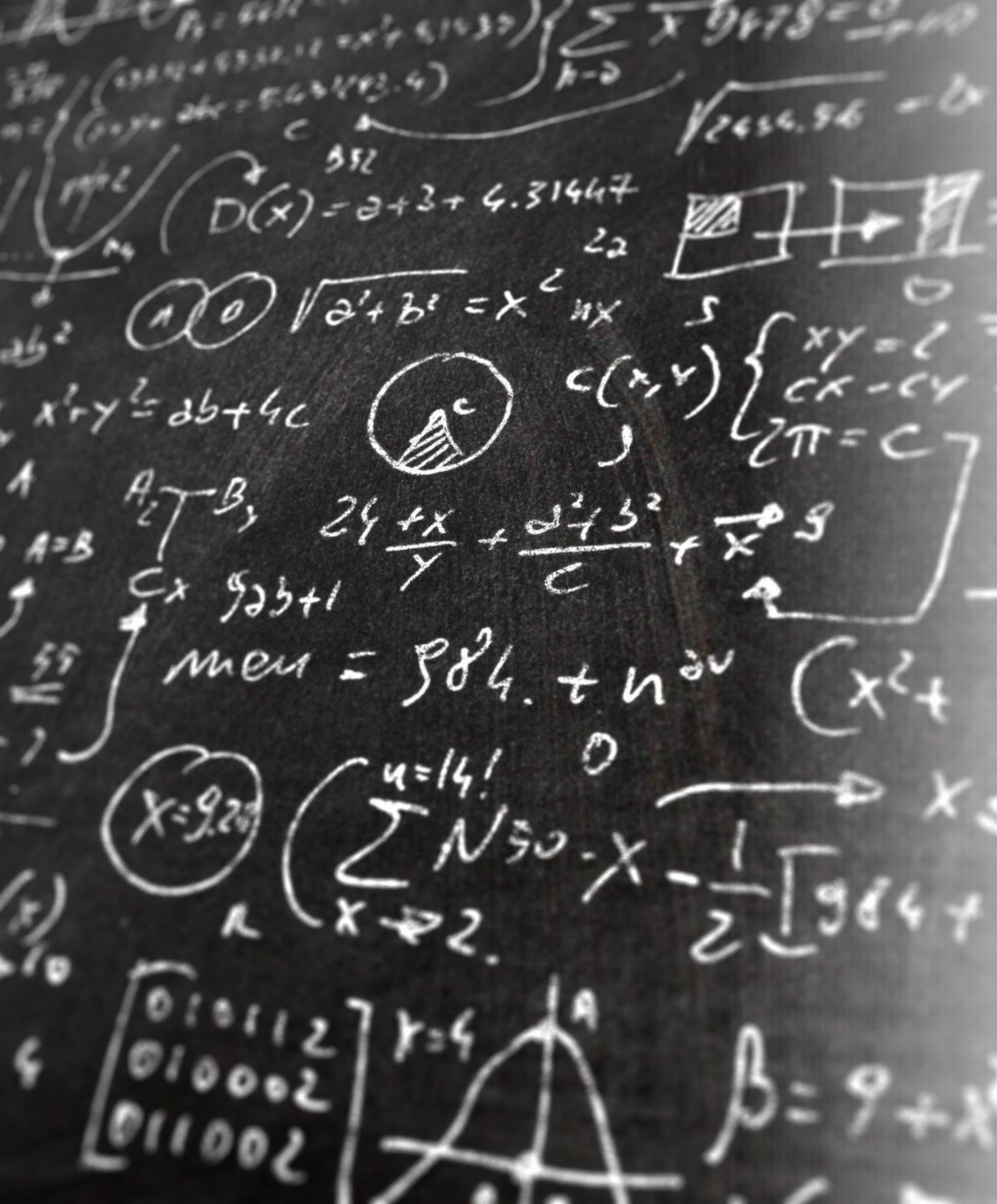
$$F = ma$$

$$F = \frac{d}{dt}mv = m\ddot{r}$$

$$v(t) = \dot{r}(t) = \dot{x}(t)\vec{i} + \dot{y}(t)\vec{j} + \dot{z}(t)\vec{z}$$

$$|v| = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$$

$$\Delta x = v\Delta t$$



2

- Equations of motion, Verlet algorithm;
- Forces (classical, AI);
- Periodic Boundary Conditions.

Newton

$$F_x = m(dv_x / dt) = m(dx^2 / dt^2) = ma_x$$

$$F_y = m(dv_y / dt) = m(dy^2 / dt^2) = ma_y$$

$$F_z = m(dv_z / dt) = m(dz^2 / dt^2) = ma_z$$

Forces

Gravitational forces

$$F = GMm/R^2 = mg$$

$$mg = m(d^2x/dt^2)$$

$$v_x = v_0 + gt$$

$$x = x_0 + v_0 t + \frac{1}{2} g t^2$$

Forces-2

Harmonic oscillator

$$-kx = m(dv_x / dt)$$

Setting k/m =1:

$$(dv_x / dt) = -x$$

Of course, there is an analytical solution to this problem. However, we are here thinking in terms of a MD code.

For ⊗t small, the time evolution of the system can be written:

$$x(t+\Delta t) = x(t) + \Delta t v_x(t)$$

Velocity at t+⊗t :

$$v_x(t+\Delta t) = v_x(t) + \Delta t a_x(t)$$
$$= v_x(t) - \Delta t x(t)$$

Kinematics
Dynamics

$$a_x = \frac{dv_x}{dt} = -x$$

Along this line, we have an iterative process to calculate (time integrate) positions and velocities, given some forces.

Numerics: improving precision

Instead of taking v(t) and v(t+⊗t), we can consider v(t+⊗t/2):

$$x(t+\Delta t) = x(t) + \Delta t v(t + \Delta t / 2)$$

$$v(t+\Delta t/2) = v(t-\Delta t/2) + \Delta t a(t)$$

$$a(t) = -x(t)$$

We need v(Δt/2), for t=t_0.

$$v(\Delta t/2) = v(0) + (\Delta t/2)a(0)$$

Example - 2 planets

Description in the plane of gravitation:

$$Fx/|F| = -x/r,$$

$$F_x = -x|F|/r = -Gmx/r^3$$

Equations to be solved:

$$m(dv_x/dt) = -GMmx/r^3,$$

$$m(dv_y/dt) = -GMmy/r^3,$$

$$r = \sqrt{x^2 + y^2}$$

Example – more (3) planets

$$m_i \frac{dv_{ix}}{dt} = \sum_{j=1,\, j \neq i}^{N} -\frac{G m_i m_j (x_i - x_j)}{r_{ij}^3},$$

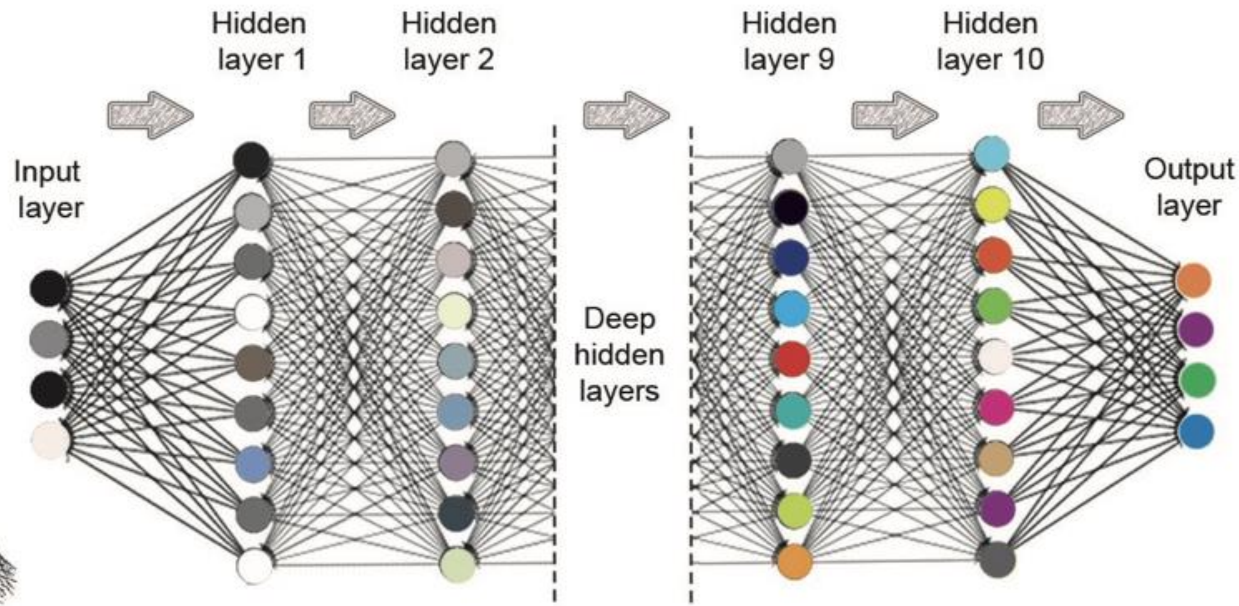
$$m_i \frac{dv_{iy}}{dt} = \sum_{j=1,\, j \neq i}^{N} -\frac{G m_i m_j (y_i - y_j)}{r_{ij}^3},$$

$$m_i \frac{dv_{iz}}{dt} = \sum_{j=1,\, j \neq i}^{N} -\frac{G m_i m_j (z_i - z_j)}{r_{ij}^3}$$

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

3-body problem

- PG Breen *et al*., „Newton vs the machine: solving the chaotic three-body problem using deep neural networks“



For n bodies…

$$\sum_i \frac{1}{2} m v_i^2 + \sum_{i \neq j} -\frac{G m_i m_j}{r_{ij}} = konst.$$

Impulse conservation

$$F_1 = -F_2$$

Action / Reaction

$$m_1 a_1 = -m_2 a_2$$

$$m_1 \frac{dv_1}{dt} = -m_2 \frac{dv_2}{dt}; mv = p$$

$$\frac{dp_1}{dt} = -\frac{dp_2}{dt}, \frac{d(p_1 + p_2)}{dt} = 0$$

$$m_1 v_1 + m_2 v_2 + m_3 v_3 + \ldots = const.$$

MD - A „Crash Course“

A simple, however quite “universal” MD program

```
program md                          simple MD program

call init                           initialization
t=0
do while (t.lt.tmax)                MD loop
   call force(f,en)                 determine the forces
   call integrate(f,en)             integrate equations of motion
   t=t+delt
   call sample                      sample averages
enddo
stop
end
```

Frenkel & Smit, p. 65

MD-Initialization

Initialization:

Prepare a *simulation box* (may correspond to a structure, a liquid, ...)

Choose a temperature, distribute velocities on the particles:

A Maxwell-Boltzmann distribution may be used at this point.

Alternative:

Velocities can be randomly assigned

The temperature is calculated from the kinetic energy, and rescaled, if necessary.

(→ A M-B distribution will be restored on time propagating)

Conservation of the linear momentum, L=0.

Velocity distributions

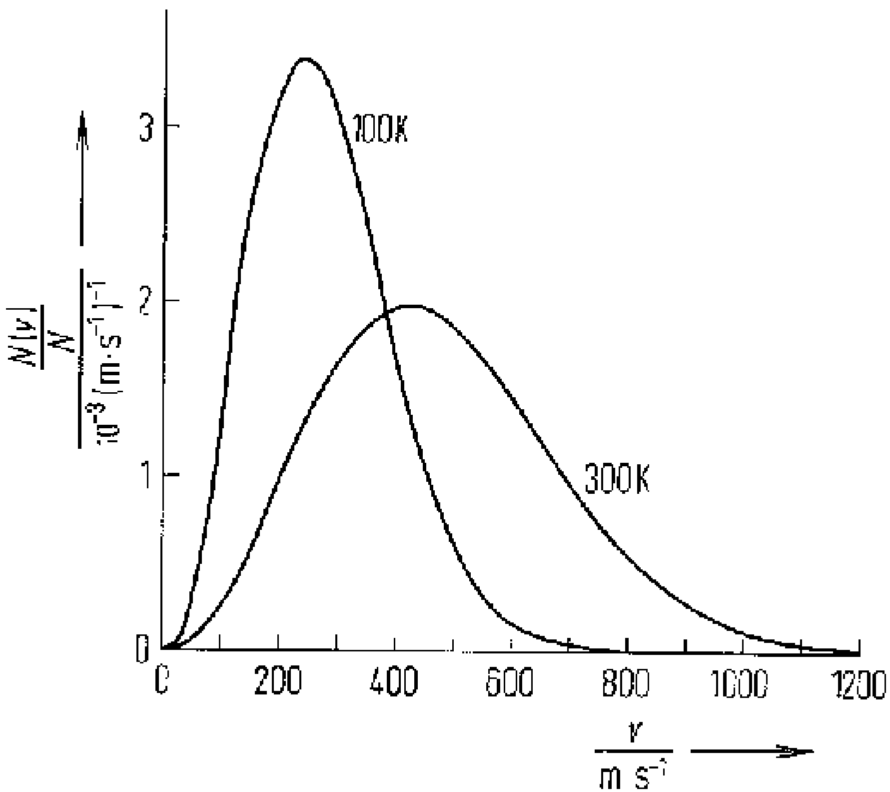


Abb. 4-16. Geschwindigkeitsverteilung für Stickstoff bei verschiedenen Temperaturen.

Maxwell-Boltzmann distribution

Initialization - Algorithm

- Loop (i) over N particles:
  - Place particles on a lattice, x(i)
  - Assign velocities, v(i)
  - vcm=vcm+v(i) (velocity of center of mass)
  - kin=kin+v(i)^2   (kin. energy)
- Done
  - set v.c.m. = 0
  - rescale velocities to *T*

MD-Equation of Motions

Verlet Algorithmus.

$$r(t+\Delta t) = r(t) + v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{3!}\dddot{r} + O(\Delta t^4)$$

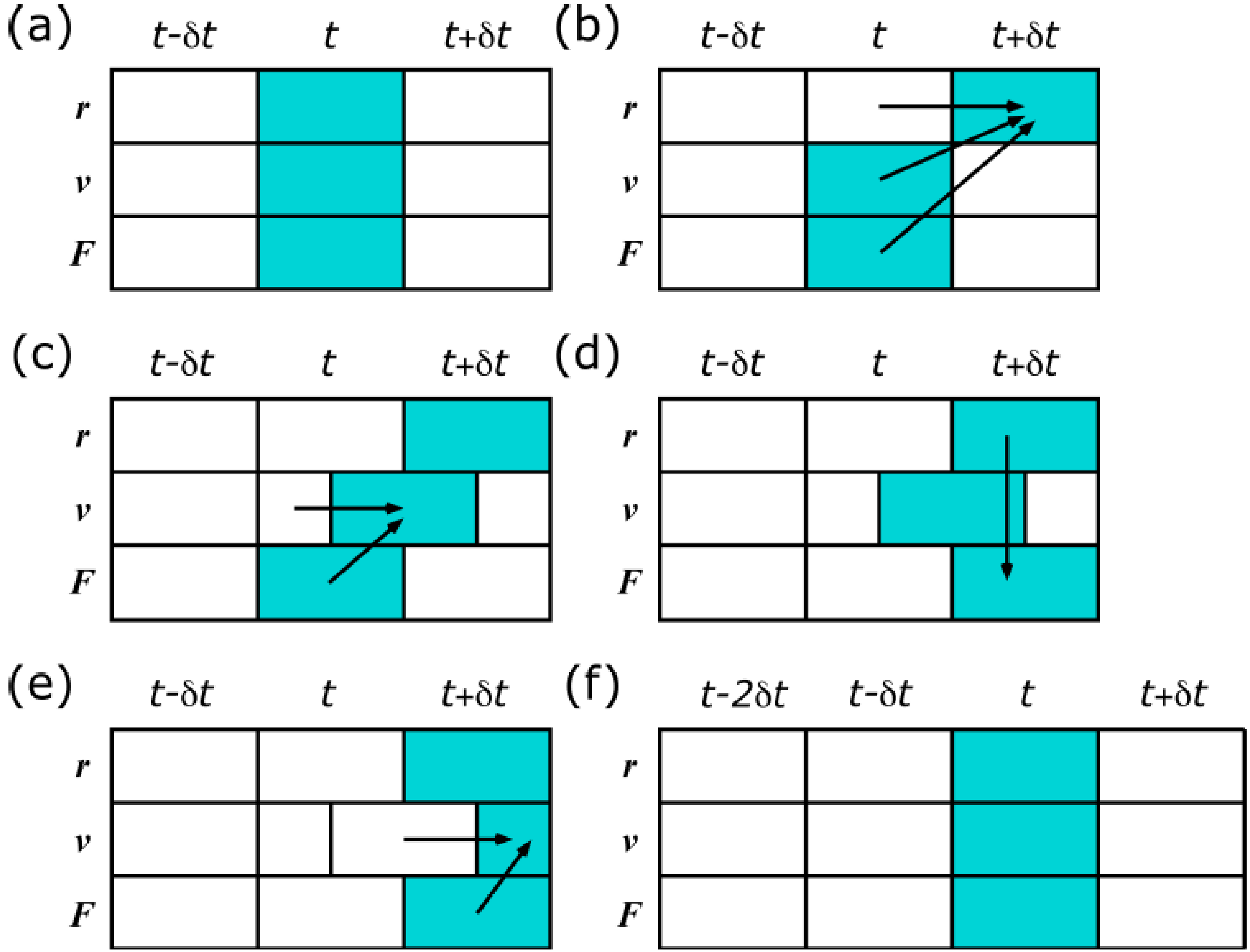
$$r(t-\Delta t) = r(t) - v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{3!}\dddot{r} + O(\Delta t^4)$$

$$r(t+\Delta t) + r(t-\Delta t) = 2r(t) + \frac{f(t)}{m}\Delta t^2 + O(\Delta t^4)$$

$$r(t+\Delta t) \approx 2r(t) - r(t-\Delta t) + \frac{f(t)}{m}\Delta t^2$$

Velocities are calculated from the position information

$$v(t) = \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} + O(\Delta t^2)$$



Verlet Algorithm - Scheme

Verlet Algorithm - Steps

Given positions, velocities & forces….
new positions can be computed ($t + \delta t$).
Velocities at $t+\delta t/2$ are computed,…
and forces at $t+\delta t$.
Velocities are computed at full step,
And the system is advanced to the next time step.

Formulas

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2$$

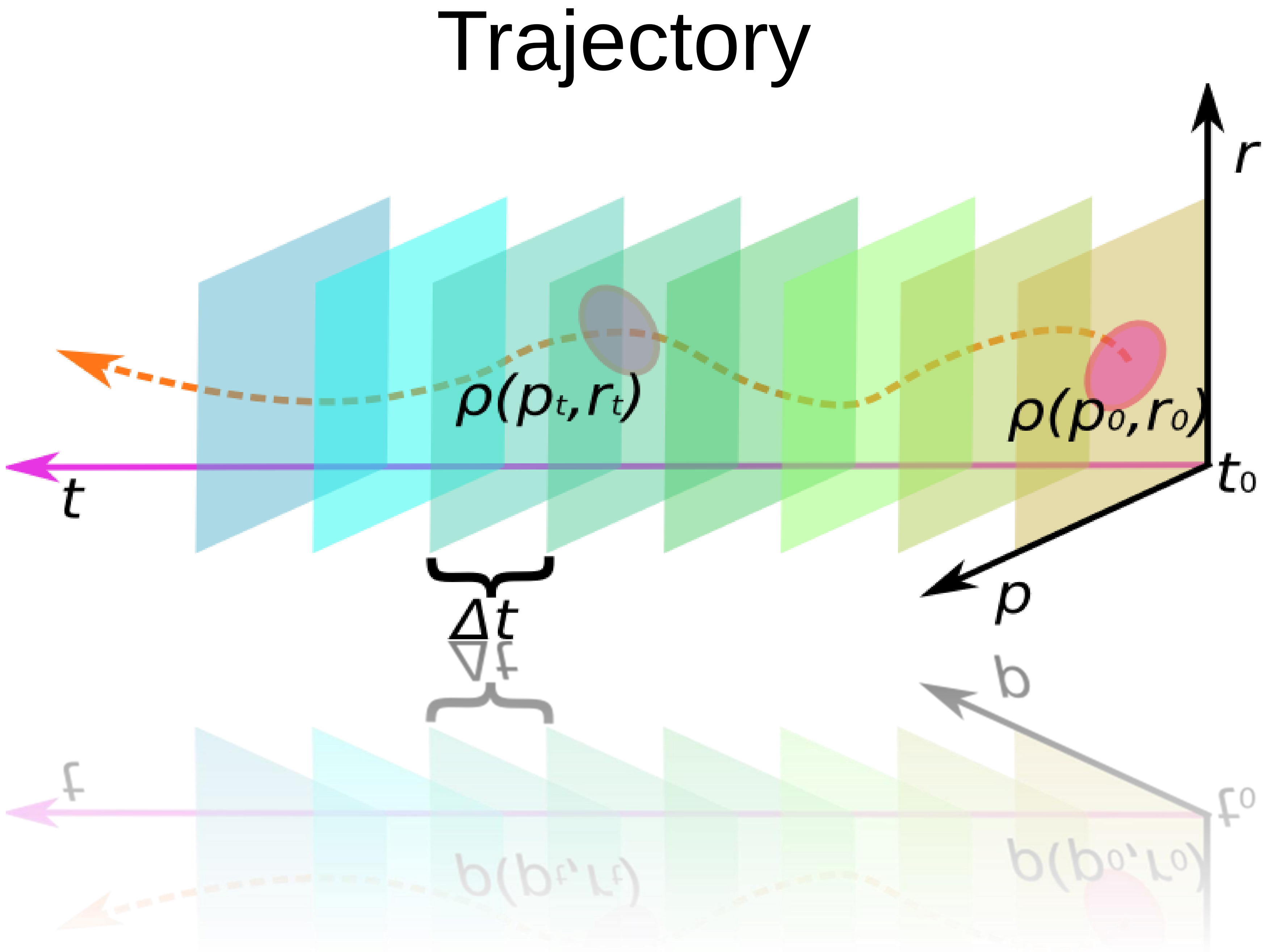
$$\mathbf{v}(t+\Delta t/2) = \mathbf{v}(t) + (1/2)\mathbf{a}(t)\Delta t$$

$$\mathbf{a}(t+\Delta t) = -(1/m)\nabla V\left(\mathbf{r}(t+\Delta t)\right)$$

$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t+\Delta t/2) + (1/2)\mathbf{a}(t+\Delta t)\Delta t$$

3

Measurements, pair/radial distribution funtion


Trajectory
r
ρ(pt,rt)
ρ(p0,r0)
t
t0
Δt
p

Measuring Properties

- Trajectories encode the time-evolution of a system. To extract properties, time-averaging on configurations is performed.

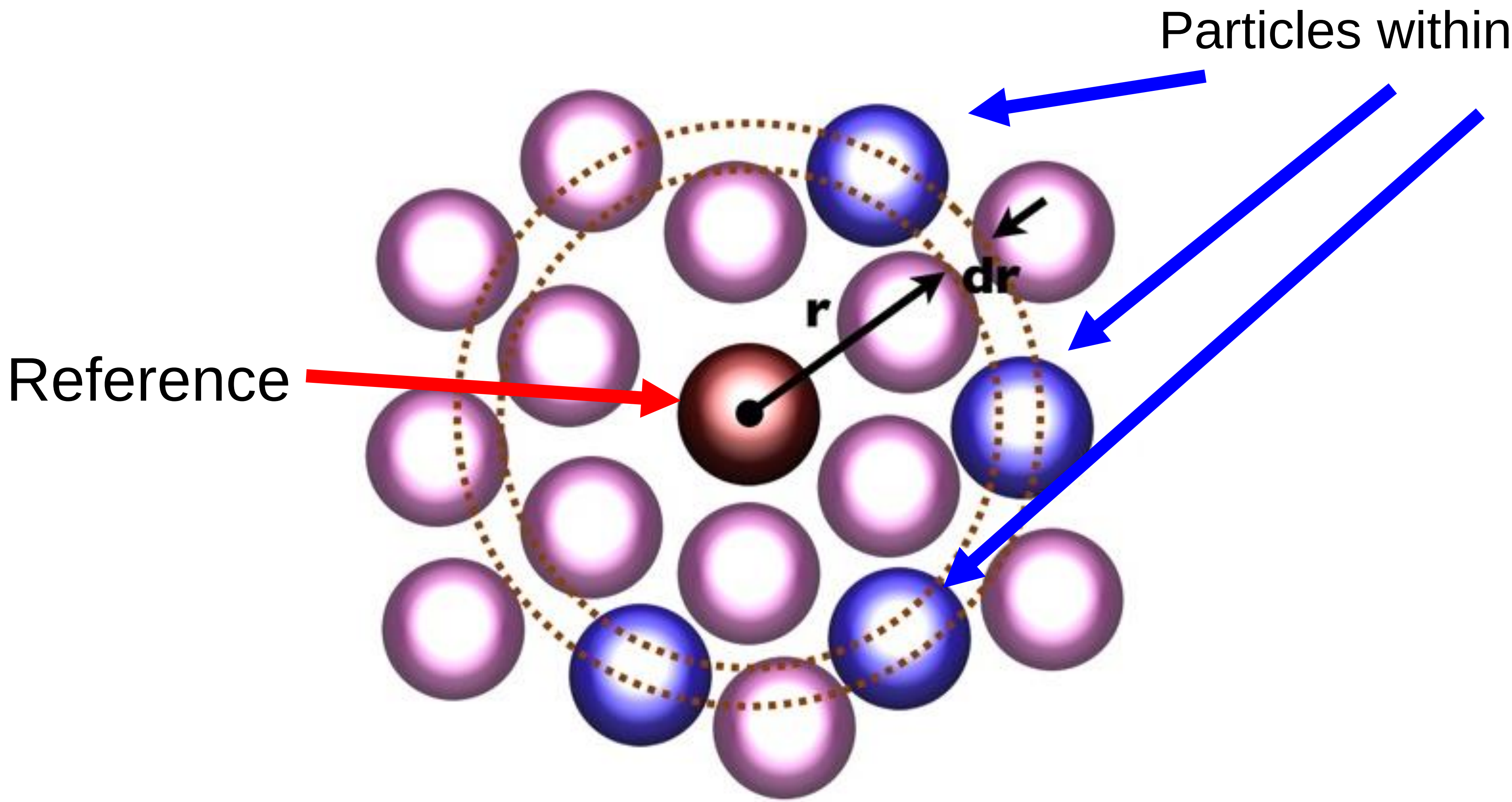
- Characterisation of:
  - Liquid/solid/gas;
  - Short/long distance correlations;
  - Typical distances (can be matched to expriments);
  - Local order/bonds.

Measures from MD: g(***r***)

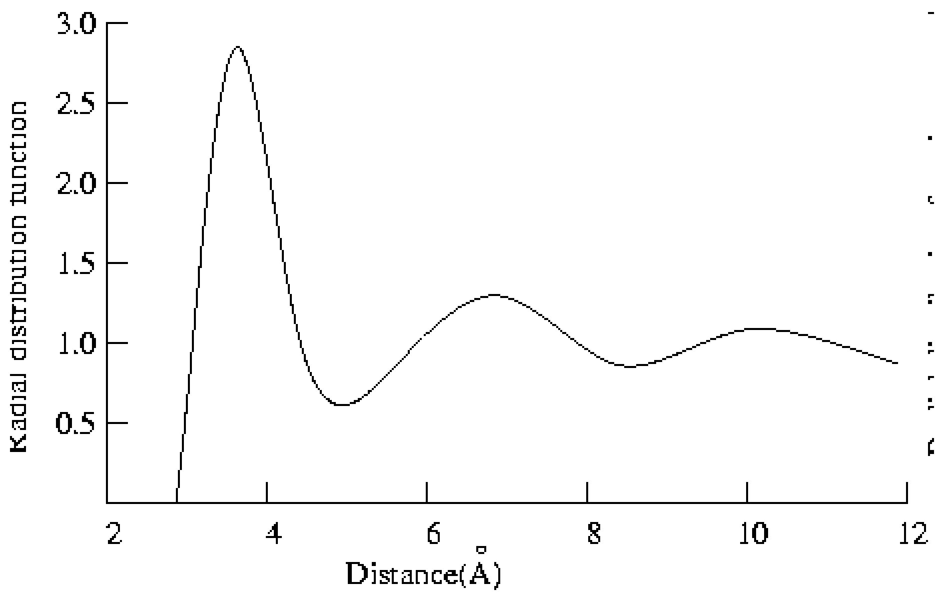
- Radial distribution function, g(***r***);
- Local structure of fluids (but also solids, amorphous systems…);
- Measurable: neutron, X-ray scattering;
- Light-scattering on colloidal suspensions;
- Central meaning in theories of liquids;
- g(***r***): ***ratio*** between the average number density ρ(r) at a distance r from any given atom and the density at a distance r from an atom in **an ideal gas** at the same overall density;
- Deviation from g(***r***) from unity reflects ***correlations*** in the system under consideration.

Number density: number of particles/volume

g(*r*) - scheme

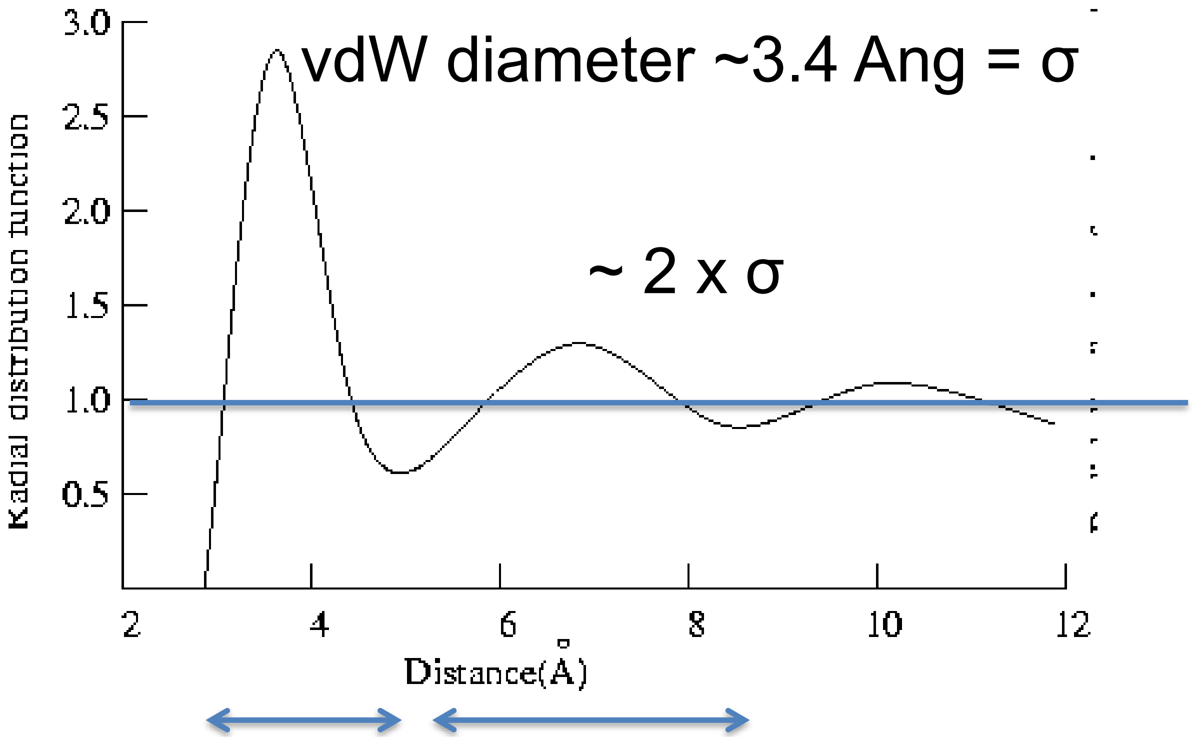


g(r): form



Liquid Argon, 100K, 100 ps MD

g(r): form



First coordination shell

Second coordination shell

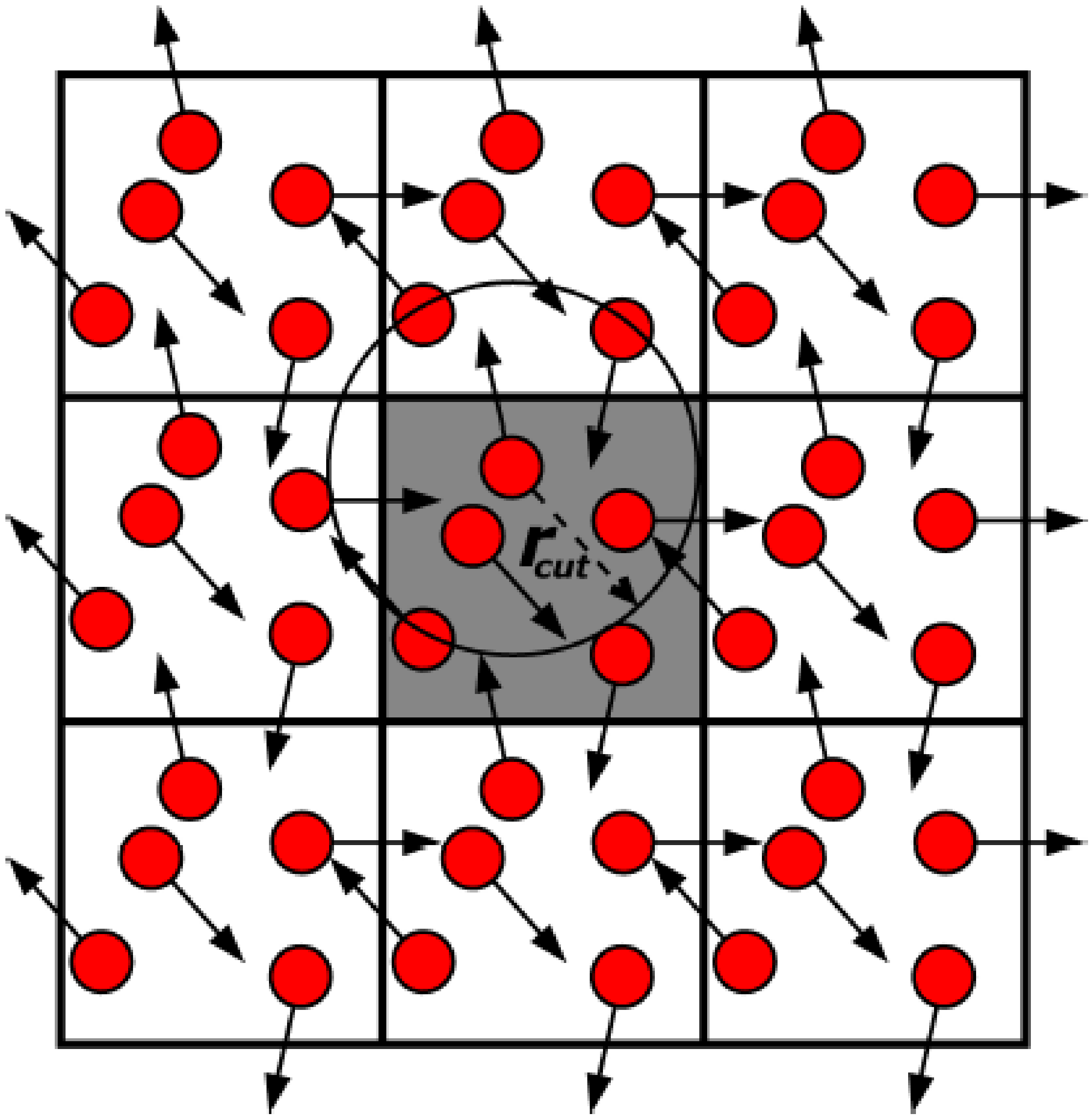
Example of nonassociated liquid, whose intermolecular structure can be understood in terms of packing.

Question

- Why are peaks equally spaced (by sigma)?
- What is the relationship to the underlying structure ?
- Similarly: what is the expected coordination number for this type of structures ?

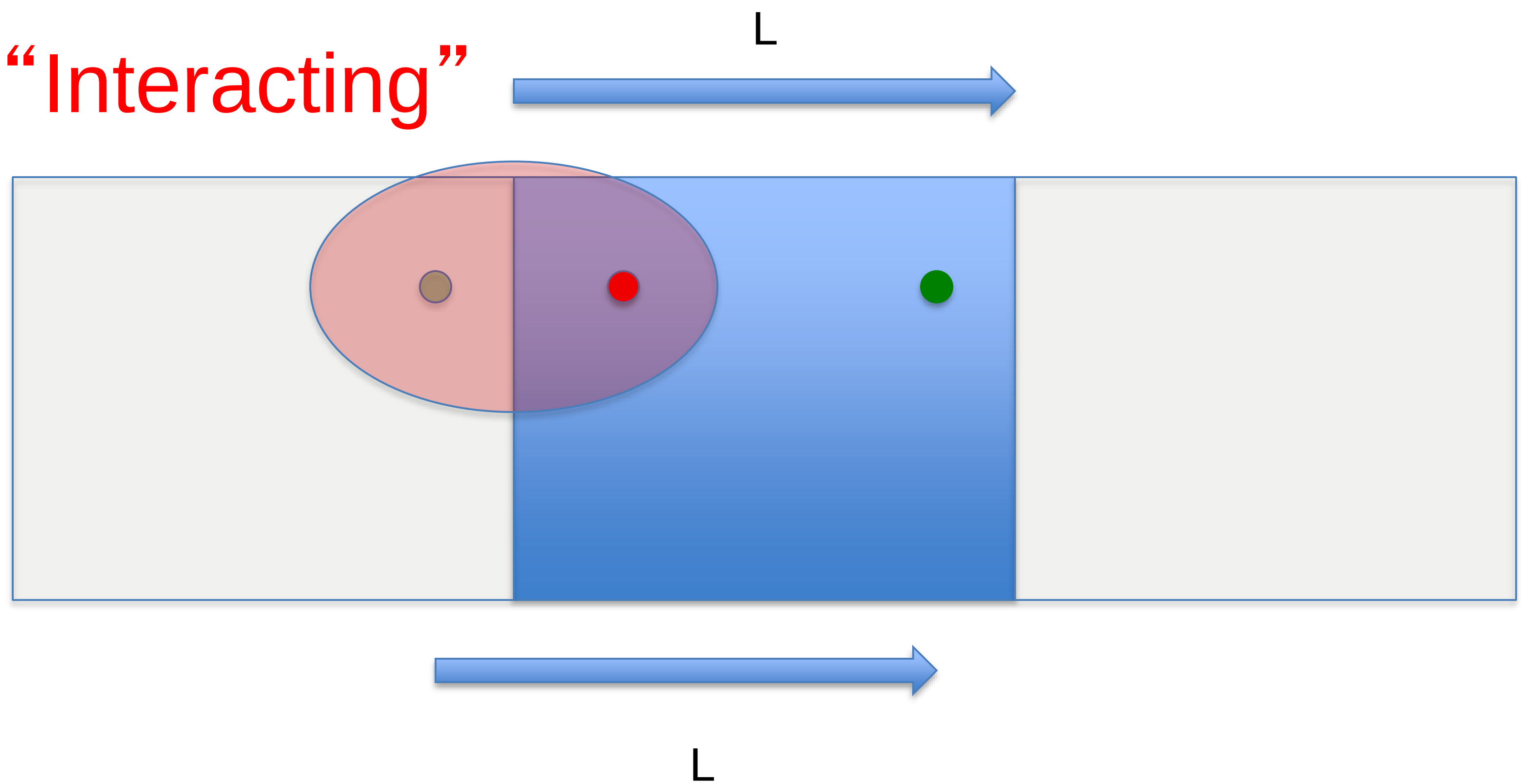
g(r): calculation

- General idea:
  - Collect distance distribution into a histogram and normalize.
- How this is done:
  - Loop over all configurations of a trajectory (time averaging);
  - Calculate all the *minimum image separations* (minimum image distances) of all pairs of atoms. Each pair contributes *2* to a histogram bin;
  - Sort these into a histogram, where each bin has a width of δr, and extends from r to $r+\delta r$;
  - Normalise.



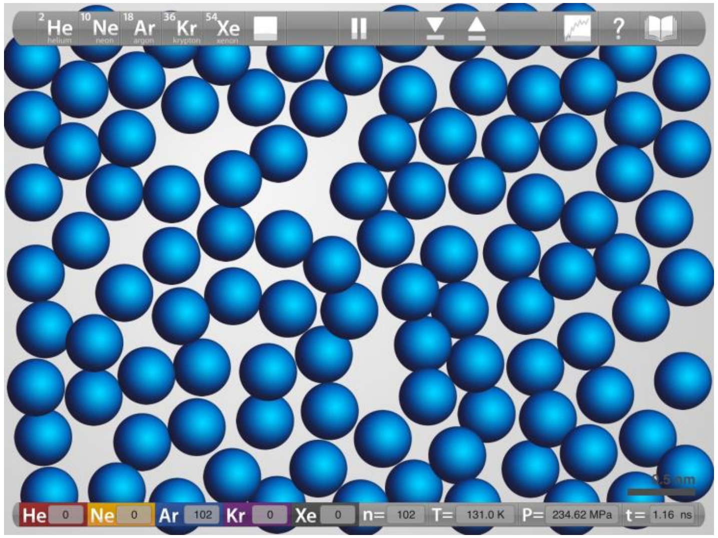
Footnote: periodic images

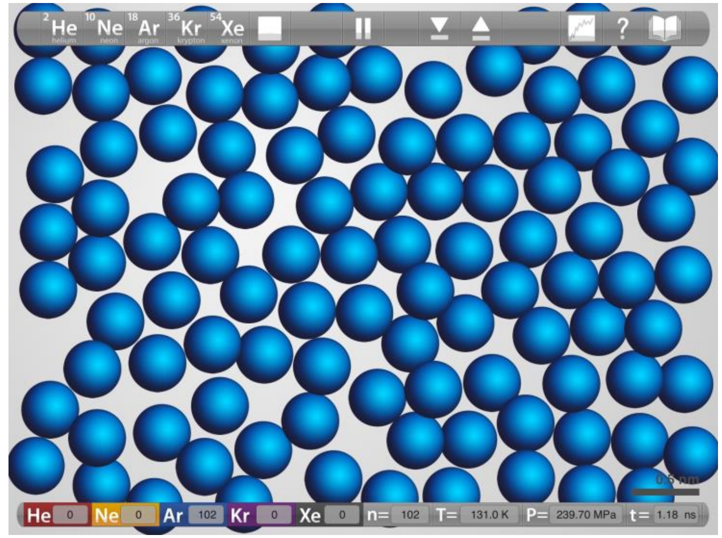
PBC & Distances

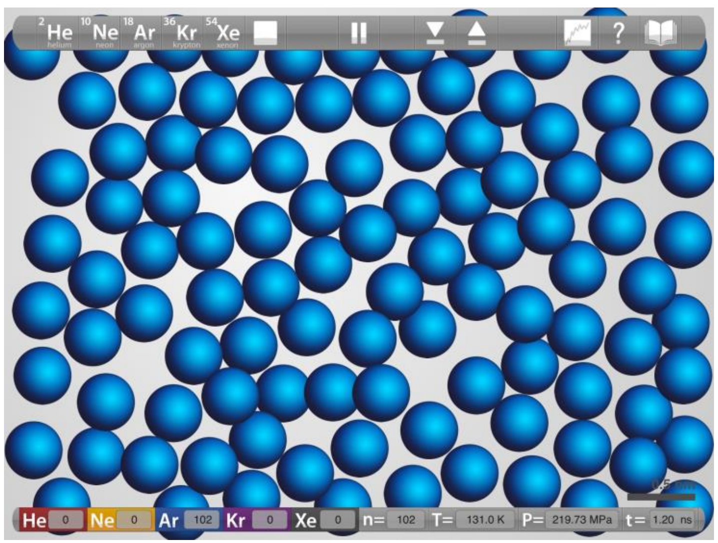


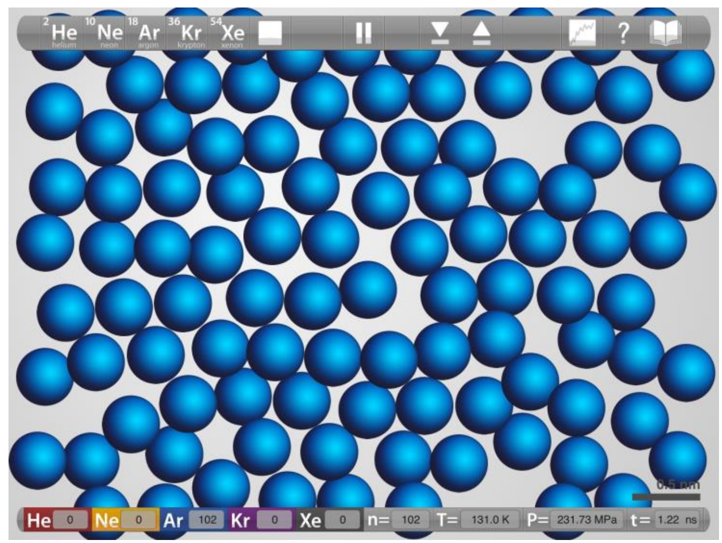
Footnote: minimum distance

Example: Ar @ 131 K

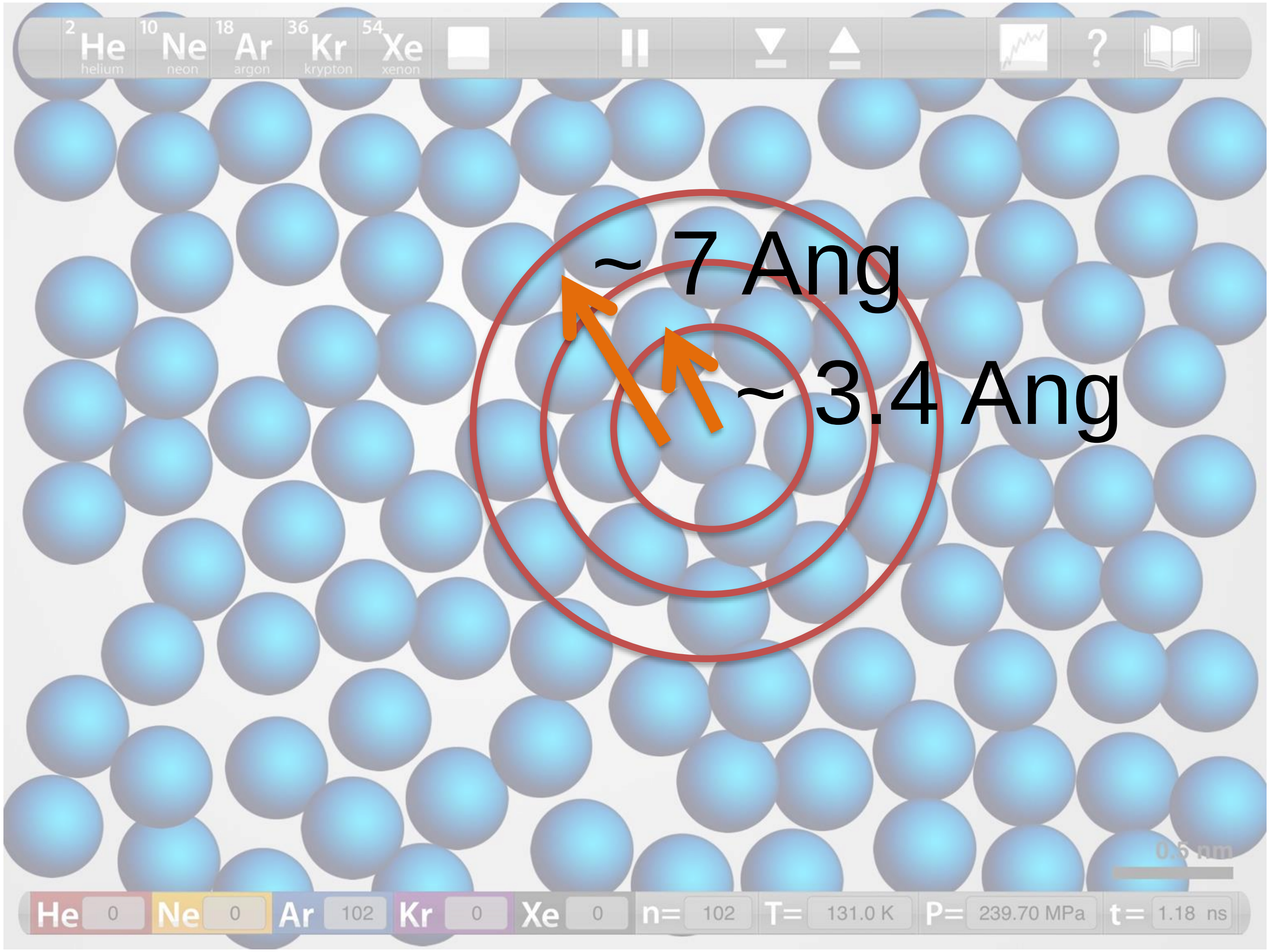




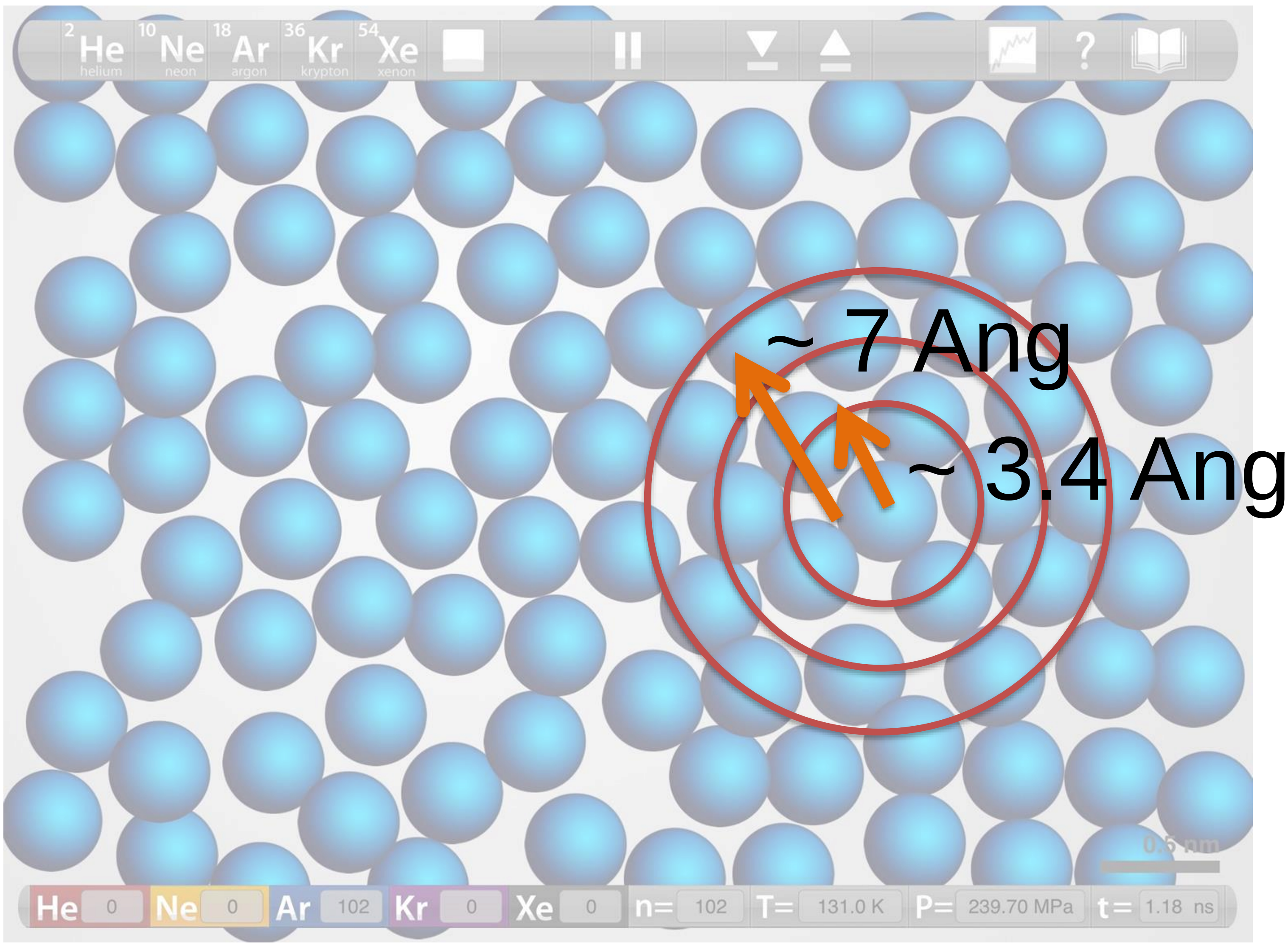




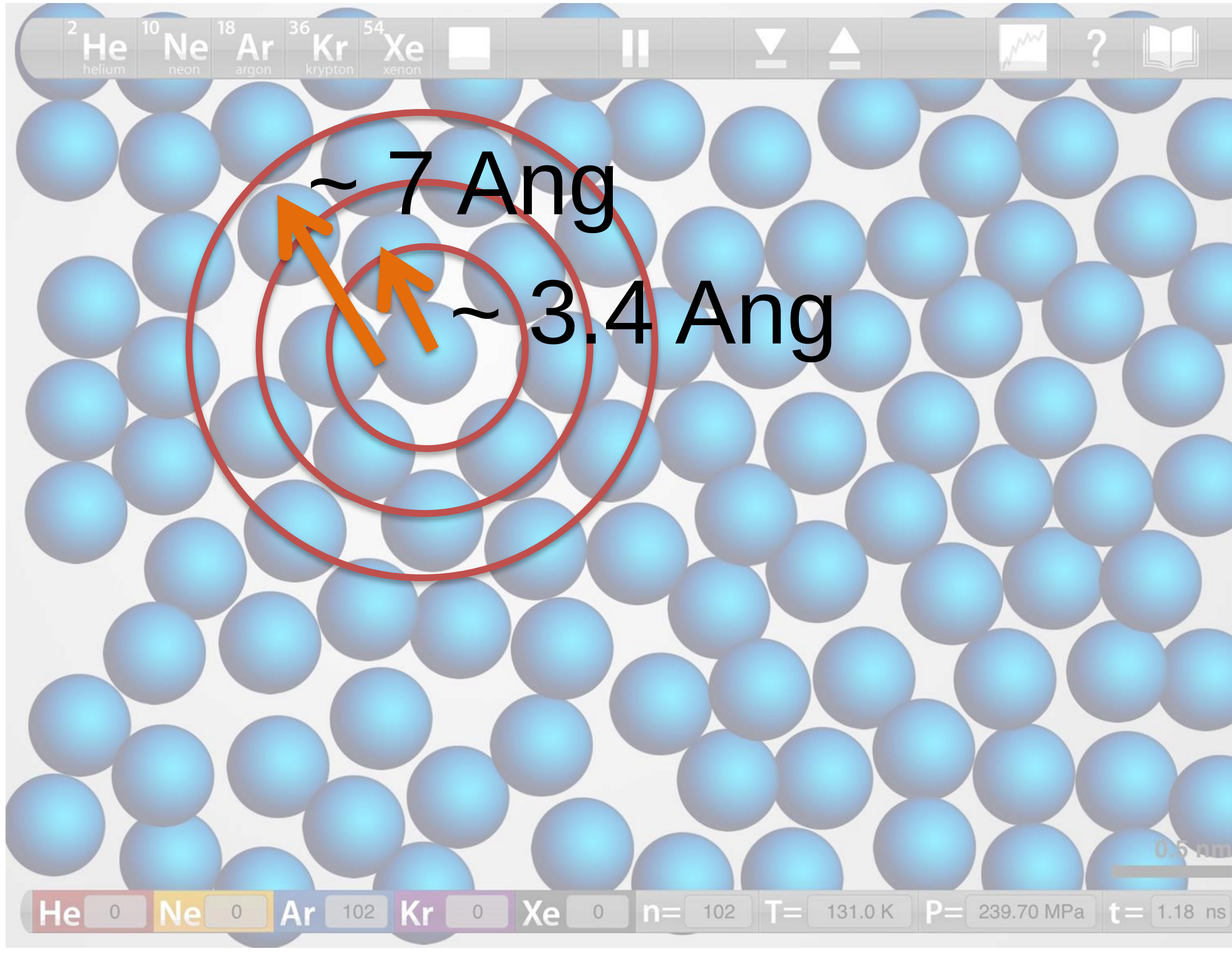
Simulation made with “AtomsInMotion” App, on a iPad

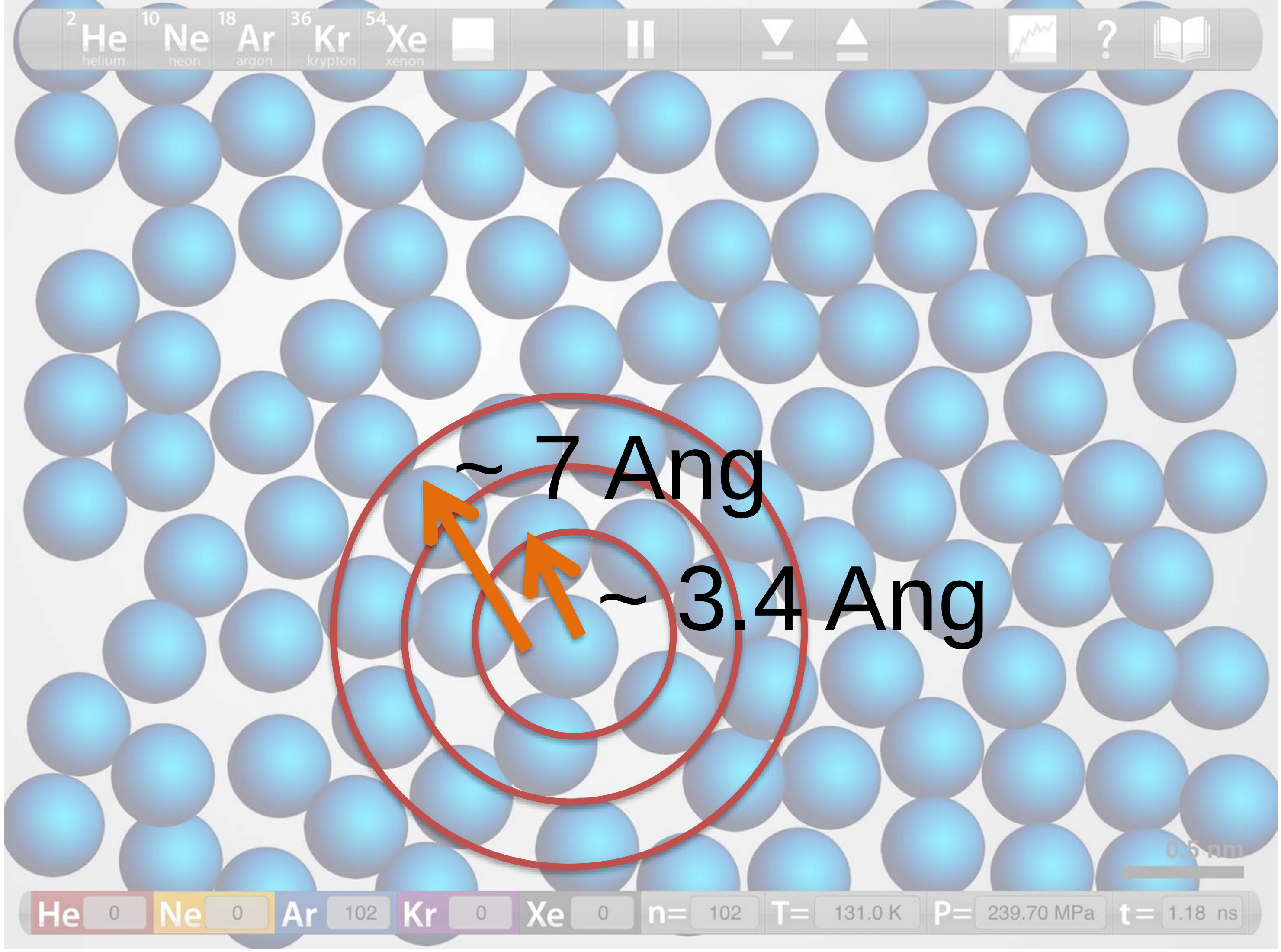


He helium
Ne neon
Ar argon
Kr krypton
Xe xenon
~ 7 Ang
~ 3.4 Ang
He 0 Ne 0 Ar 102 Kr 0 Xe 0 n= 102 T= 131.0 K P= 239.70 MPa t= 1.18 ns



~ 7 Ang
~ 3.4 Ang
2 He helium
10 Ne neon
18 Ar argon
36 Kr krypton
54 Xe xenon
He 0
Ne 0
Ar 102
Kr 0
Xe 0
n= 102
T= 131.0 K
P= 239.70 MPa
t= 1.18 ns


~ 7 Ang
~ 3.4 Ang
2 He helium
10 Ne neon
18 Ar argon
36 Kr krypton
54 Xe xenon
He 0
Ne 0
Ar 102
Kr 0
Xe 0
n= 102
T= 131.0 K
P= 239.70 MPa
t= 1.18 ns


~ 7 Ang
~ 3.4 Ang
2 He helium
10 Ne neon
18 Ar argon
36 Kr krypton
54 Xe xenon
He 0
Ne 0
Ar 102
Kr 0
Xe 0
n= 102
T= 131.0 K
P= 239.70 MPa
t= 1.18 ns

Normalisation

- Given a bin in the interval (r, r+δr), containing n_{HIS} pairs, then the *average* number whose distance lies in this interval is:

$$n_{AVERAGE} = n_{HIS} / (N' \; t_{RUN})$$

- *N* is the number of atoms, τ_{RUN} is the number of history steps (multiplied by Δt).

Normalisation -2

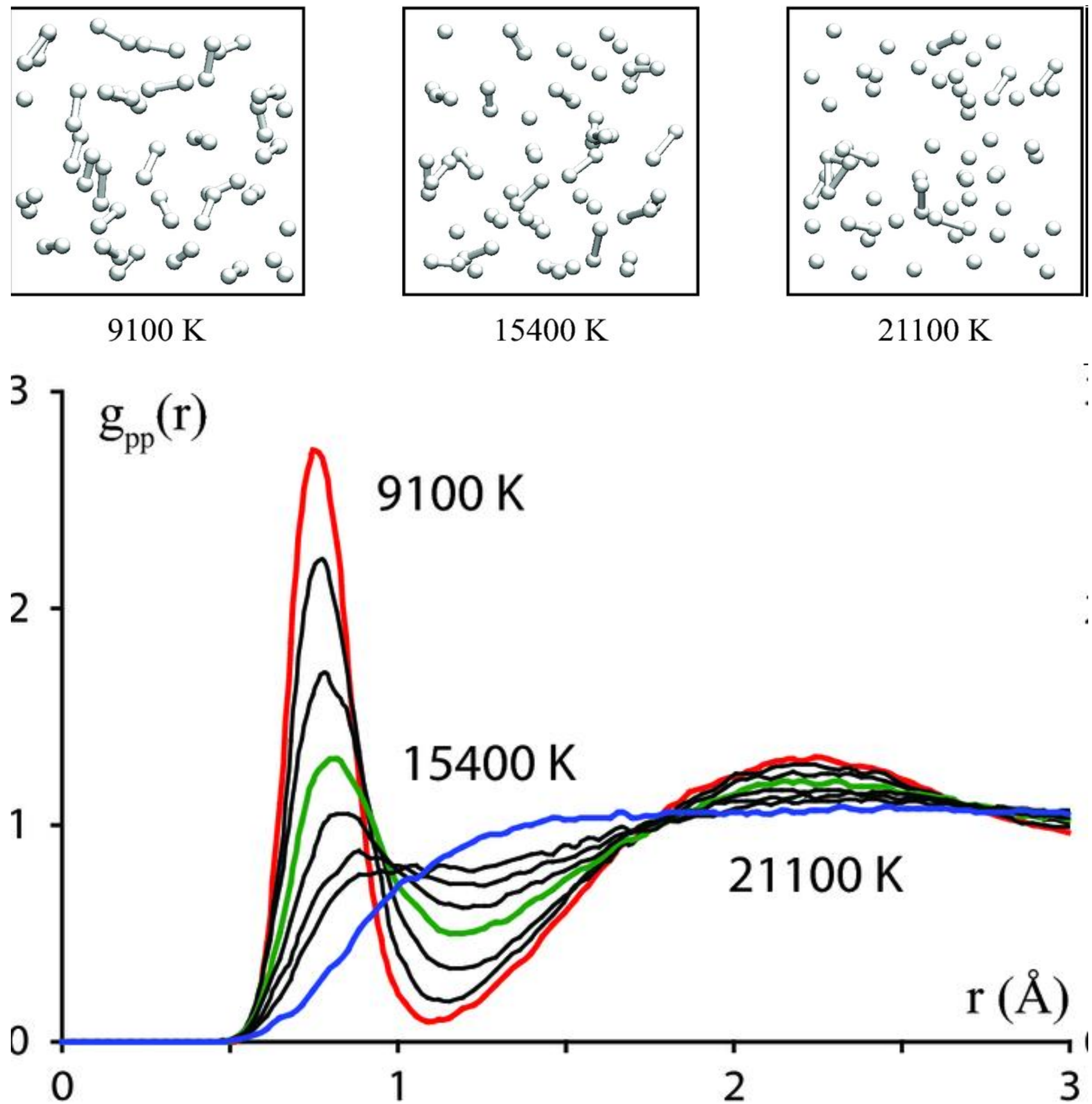
- The average number of atoms in the same interval in an *ideal* gas at the same density ρ is:

$$n^{ideal} = \frac{4pr}{3}\left[(r + dr)^3 - r^3\right]$$

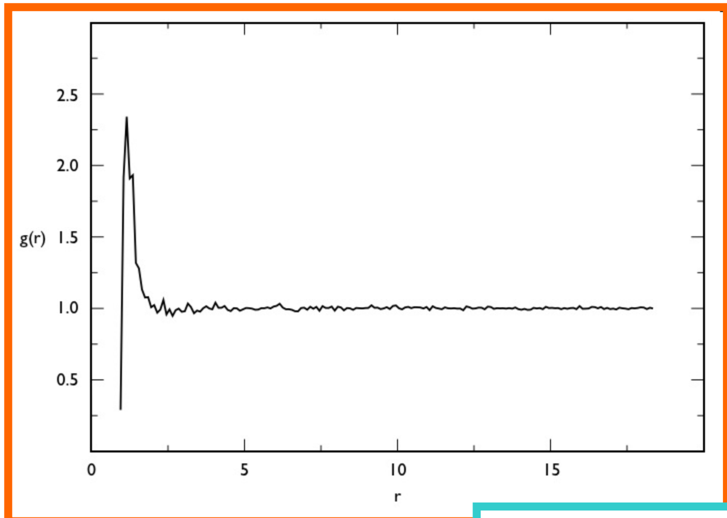
- By definition, the radial distribution function is:

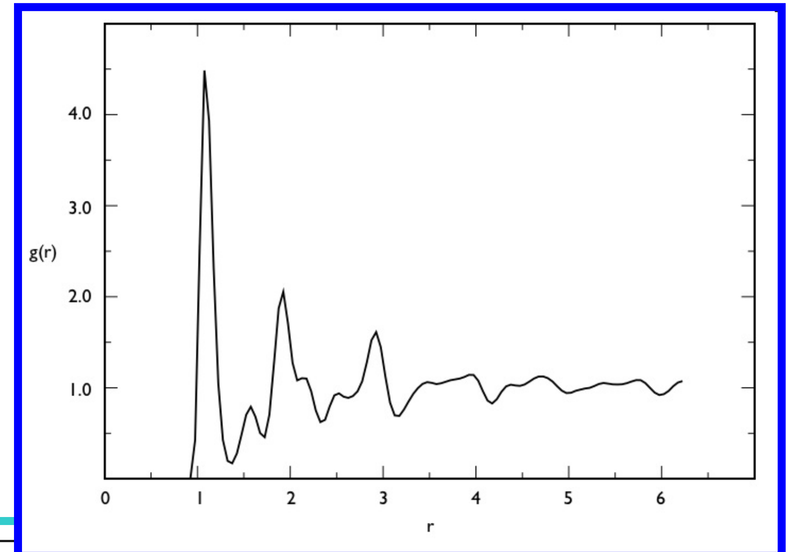
$$g(r + \frac{1}{2}dr) = n_{AVERAGE} / n^{ideal}$$

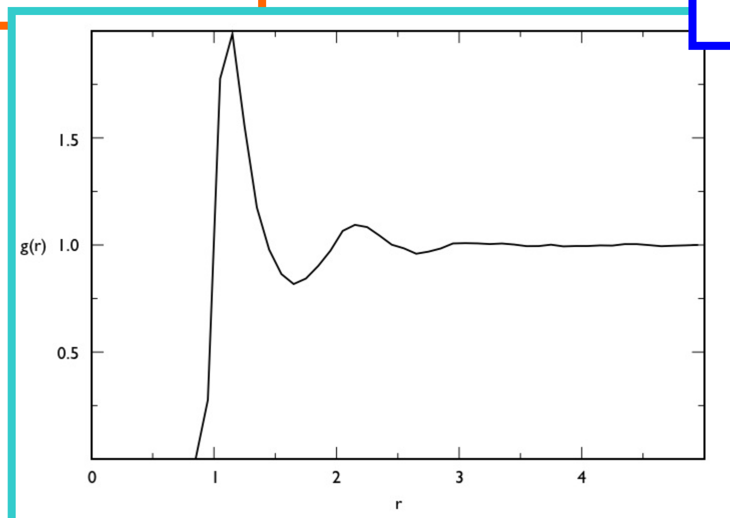
g(r) form: Hydrogen



Gas/liquid/solid







http://www.matdl.org/matdlwiki/index.php/softmatter:Radial_Distribution_Function

Questions

- When do we expect g(r) to deviate from 1 ?
- At larger distances, what is the expected value of g(r) ? (Hint: look at the definition, or reason on correlation/interaction between particle positions).
- Can we get information about particle mobility from g(r) ? (Hint: think about peak separation – what does it mean if peaks are fully separated, what does it imply if there is partial/substantial overlap?).

Strong intermolecular interaction

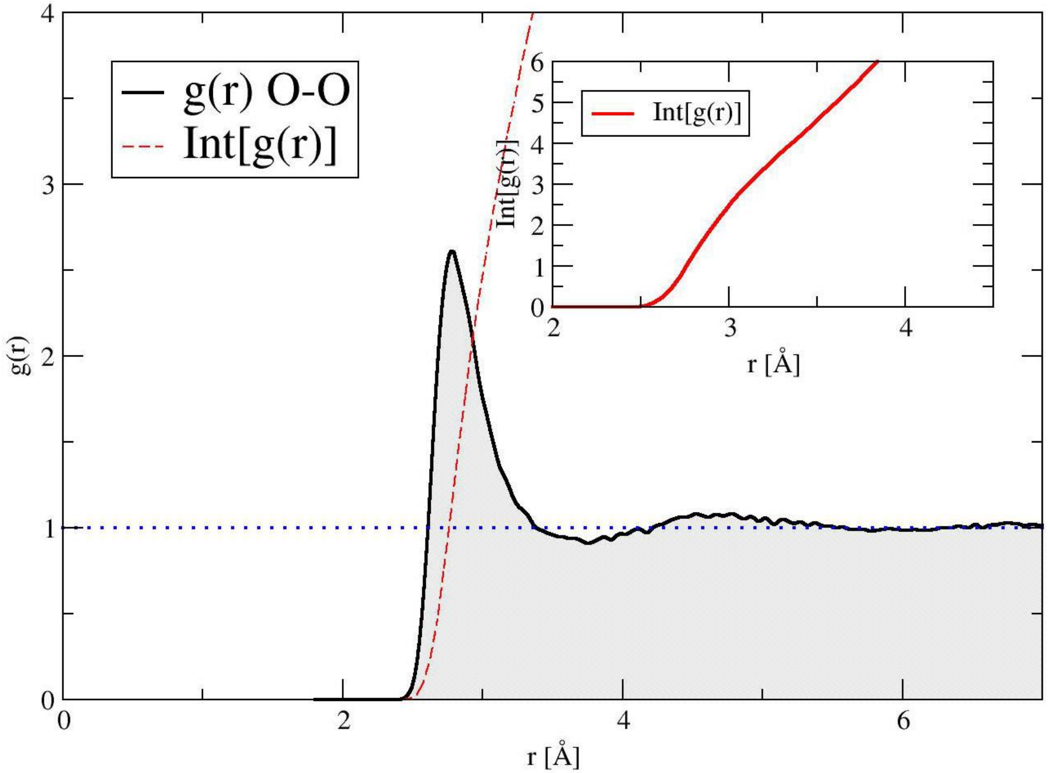
Depending on the type of intermolecular interaction, g(r) may contain additional features.

In the presence of hydrogen bonds, for example, mutual molecular orientations are rather fixed, such that one characteristic distance can be expected.

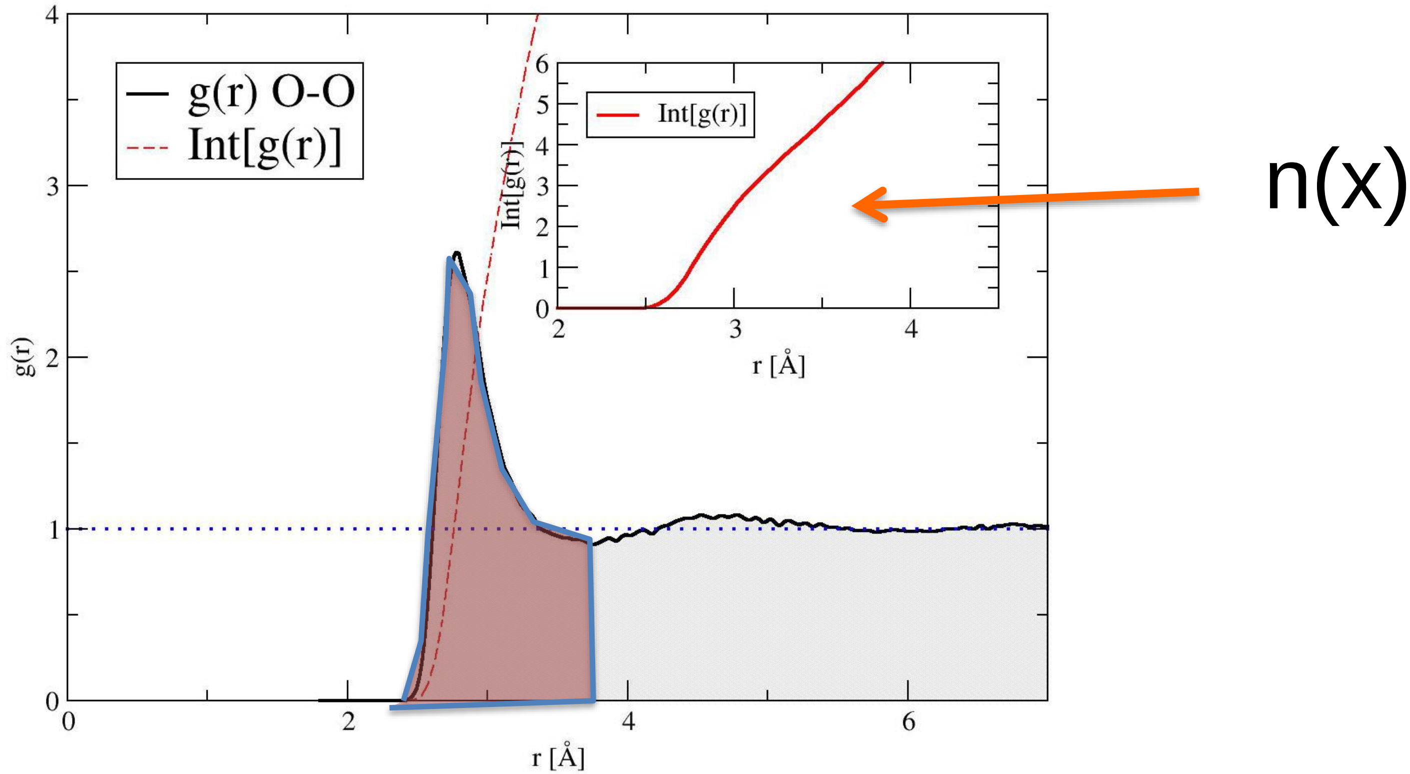
On the next slide, the first peak corresponds to O-O vdW contacts, associated with O-H….O bonds. Structural details of water can be found on subsequent slides.

Water orientation is dictated by H-bonds, the structure of solid water, ice, is not completely washed out in liquid water.

g(x) of water (liquid)



Coordination number from g(x)

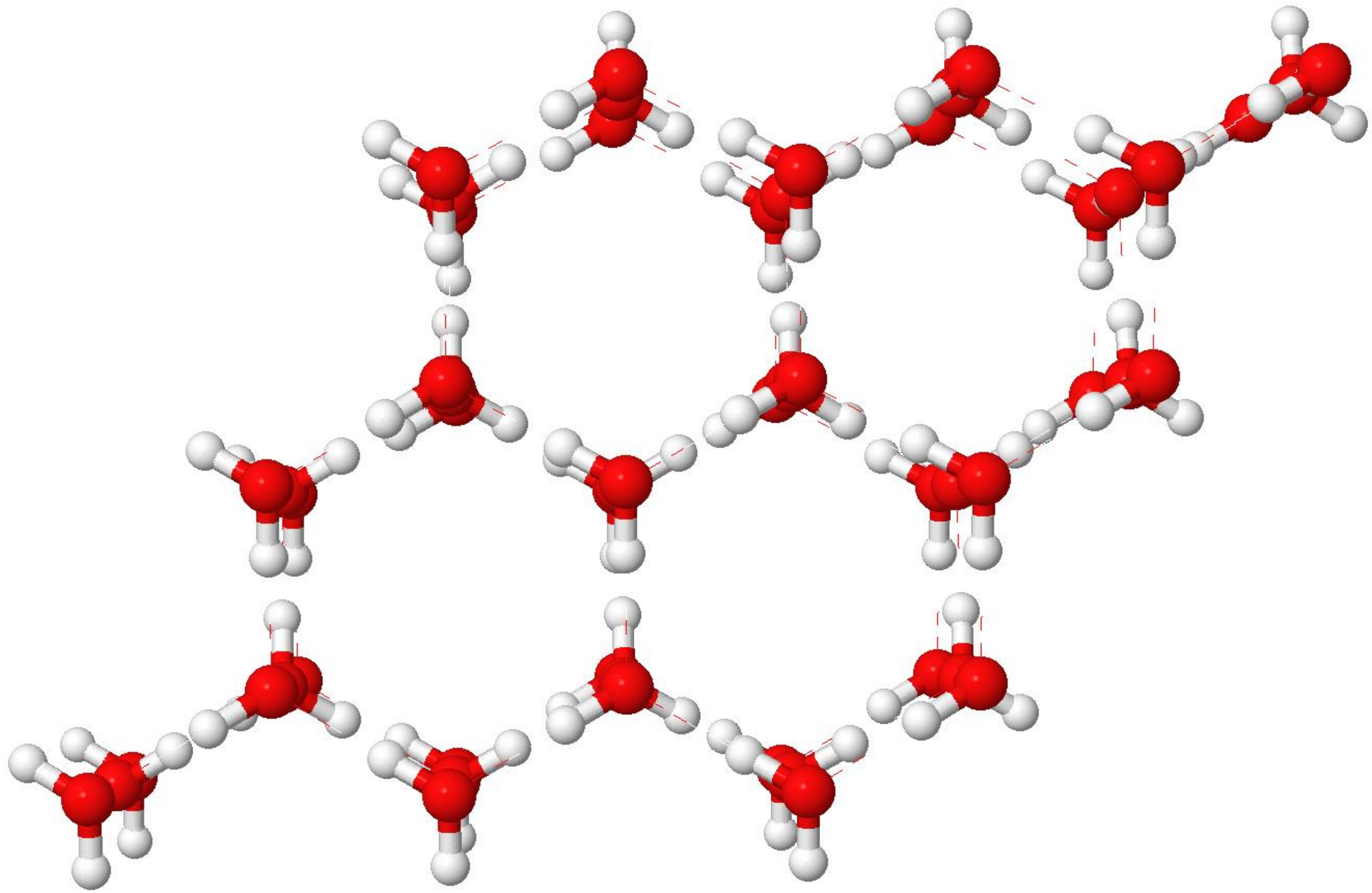


$$n(x) = 4\pi\rho \int_0^r x^2 g(x)\,dx$$

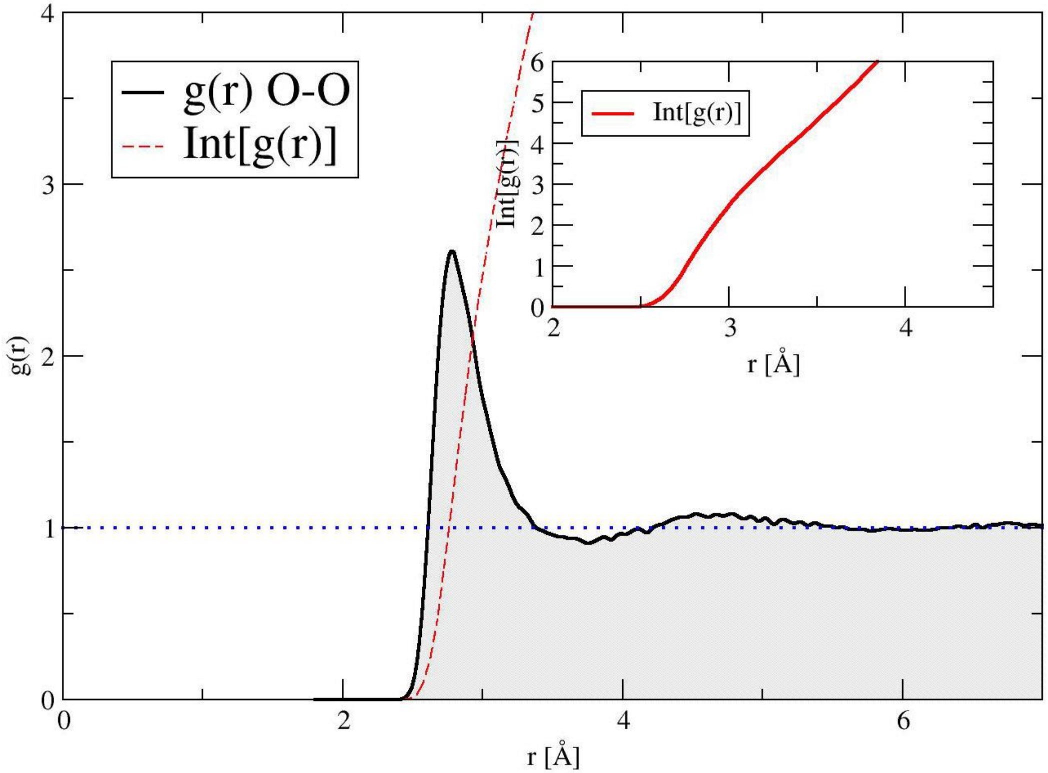
Number of neighbors within a distance r from a central atom = Coordination sphere

For liquid water, this integrates to n(x) > 4.

Crystal structure of water



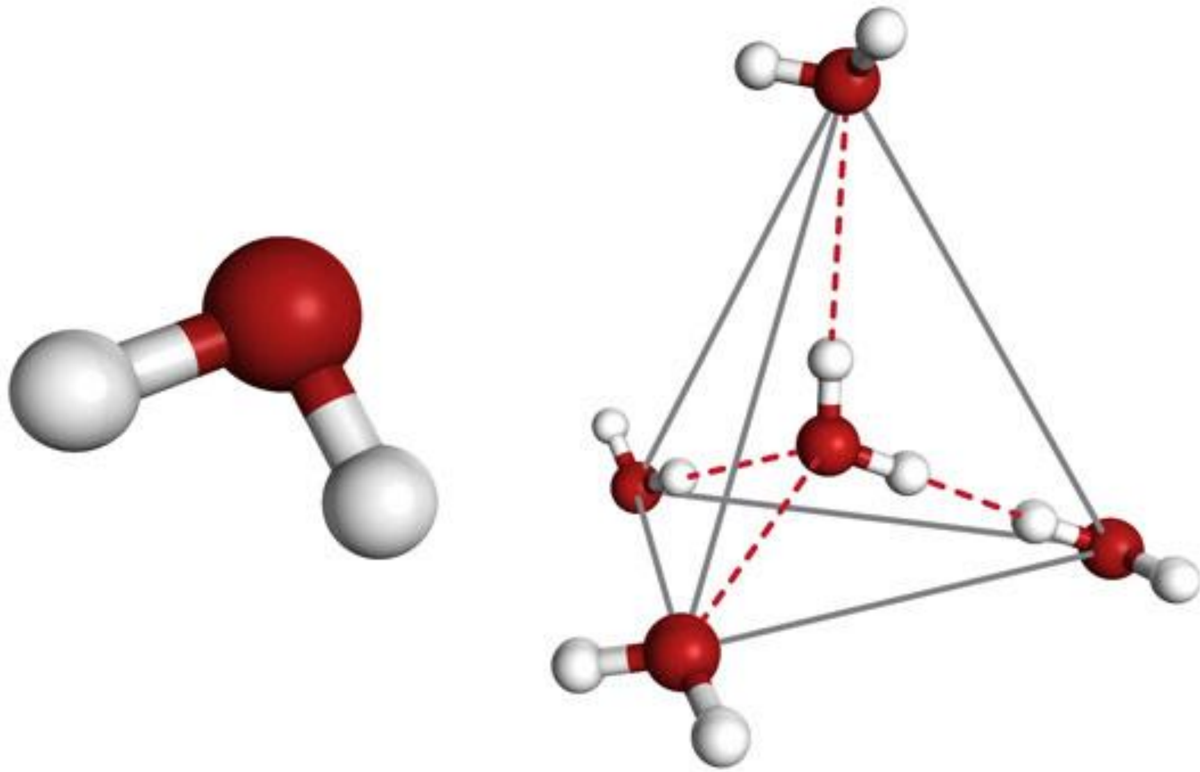
First/Second Peak



~2.8 Å ~4.5 Å = 1.6 x 2.8 Å

Question: where does the factor 1.6 come from ?

Ice: local structure



Example of **associated** liquid, in which local order is invoked by bonds, resulting in structures different from packing-controlled (nonassociated) liquids.

http://scitechdaily.com/first-structural-observations-liquid-water-temperatures-minus-51-degrees-fahrenheit/

Questions:

- Based on local tetrahedral arrangement, where can the mean peak for g(r) [H-H] and g(r) [O-H] be expected ?
- Where would the peak for intramolecular g(r) [H-H] be expected?
- Is the water structure a dense structure ?
- What are the g(r) features that point to a liquid ?
- On imposing pressure (at constant temperature), which will disrupt hydrogen bonds, do we expect the order (entropy) of water to decrease or increase ?

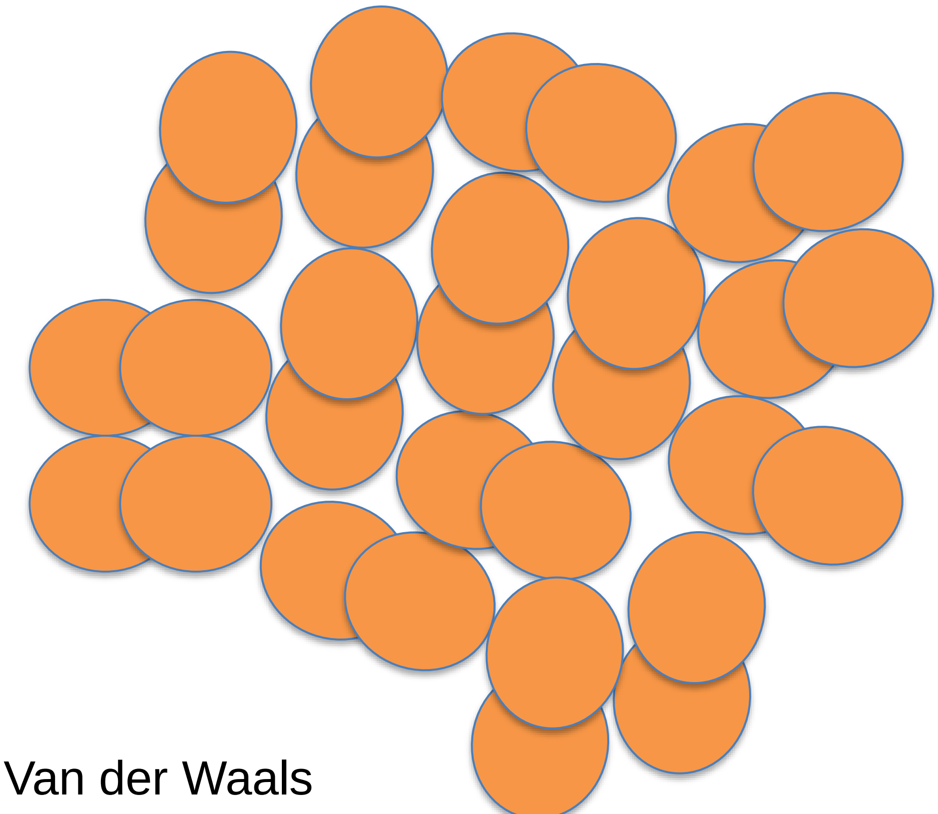
Weak(er) intermolecular interaction

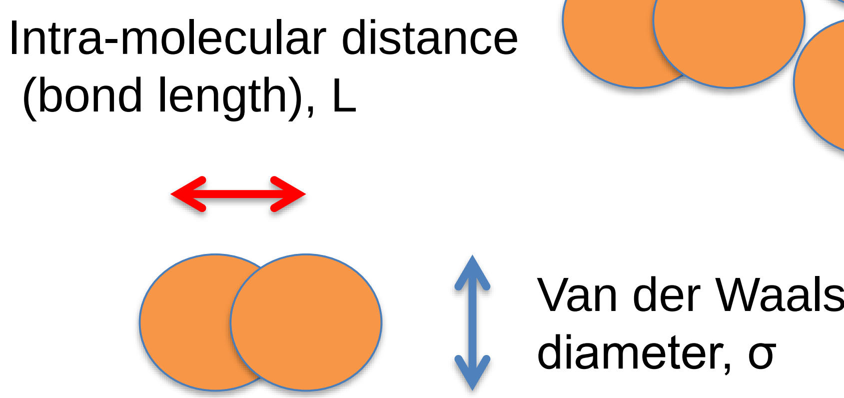
For weaker (non H-bonded) intermolecular interactions, the mutual orientation of molecules is broader, and (in liquid/solids) dictated by the packing & by the alignment of molecules.

The g(r) will contain a first (main) peak, corresponding to the van der Waals diameter of the molecule of interest, plus a second peak (or shoulder), which results from the fact that molecules touch each other, and their orientation is not so restricted as in a situation of H-bonded molecules. A broad area is normally delimited by those two peaks (peak overlap).

This is the case for diatomic molecules (liquid elements for example), or hydrocarbons.

Structure of a liquid with diatomic molecules (packing of peanuts)

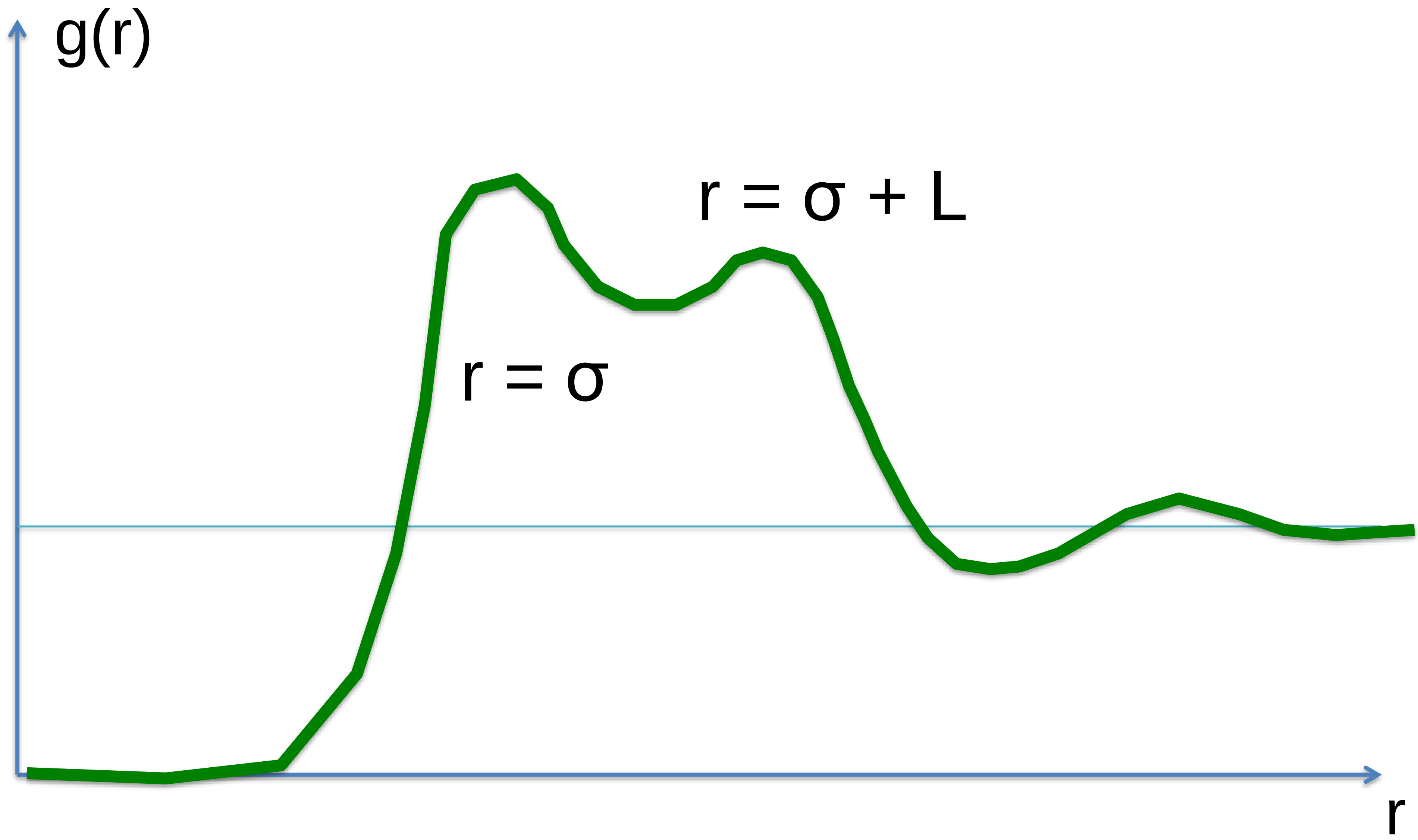




Question

- Predict the g(r) of liquid N_2, including inter- and intramolecular distances.

g(r) of ethane/propane



There are two lengths in the system, sigma and sigma + L.
Sigma is sharper and better defined, on the contrary, due to a more or less random orientation of molecules, there is a range of intermolecular distances between sigma and sigma + L.

g(r) appears therefore less rich of features and the orientation randomization washes out any finer structuring.

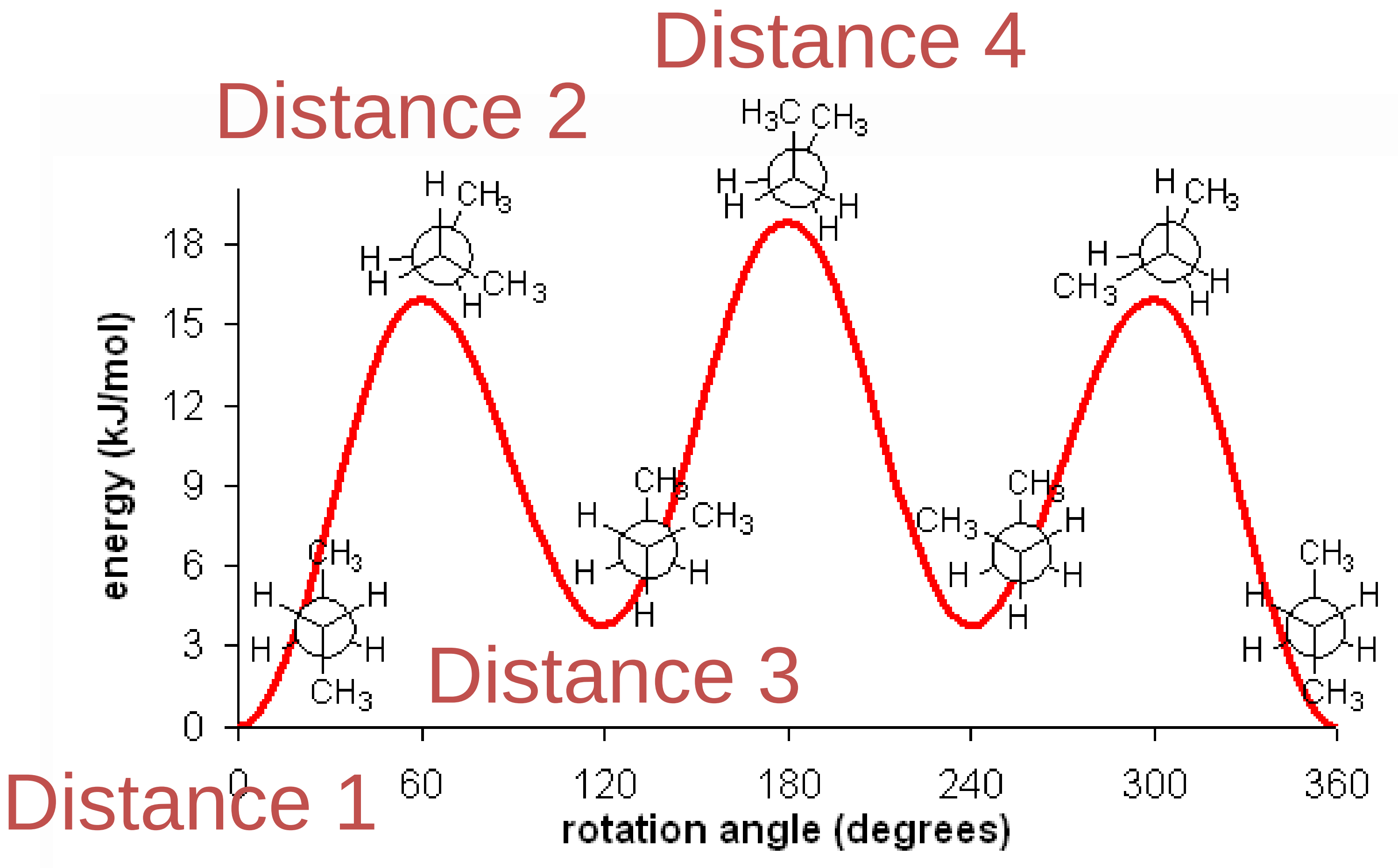
The role of conformations

Conformations may affect g(r).
For hydrocarbons, for example, $g(r)_{C\text{-}C}$ will show a first sharp peak, corresponding to a C-C bond distance.

A second peak is expected, corresponding to the distance between two terminal C atoms in C-C-C. There is only one distance for them.

Additionally, depending on the statistical weight of a conformation, additional peaks, more or less sharp, may appear. They result form C-C-C-C. Clearly, depending on the conformations, more than one peak can be expected.

Conformations of butane



Questions

How many distances ?

Which ones will be visible ?

How will different energies be reflected into the g(r) ?

Sketch the expected g(r) [C-C],
focusing on intramolecular interactions.

Lyapunov Instability

$$r(t) = f[r^N(0), p^N(0); t]$$

$$r'(t) = f[r^N(0), p^N(0) + e; t]$$

Small perturbation ε on the momenta

Effects of the instability

$$\left|\Delta r(t)\right| = \left|r(t) - r'(t)\right| \propto \varepsilon \cdot \exp(\lambda t)$$

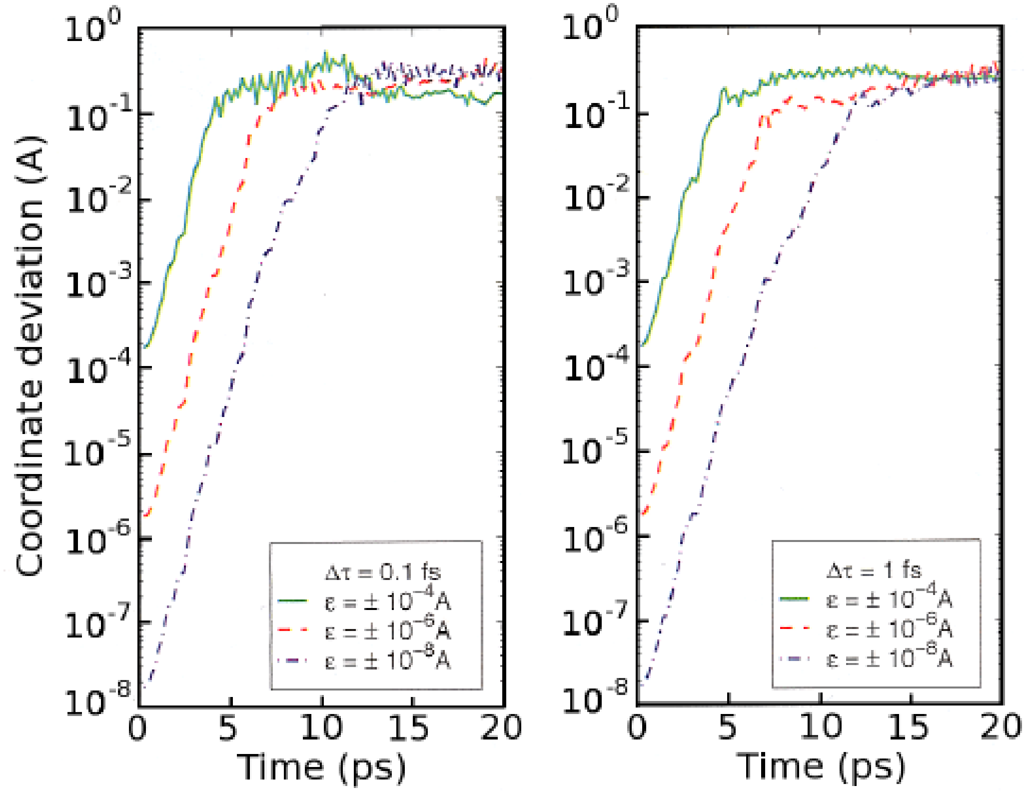
Long-time exponential divergence of initially close trajectories

$$\varepsilon \propto \Delta_{max} \exp(-\lambda t_{max})$$

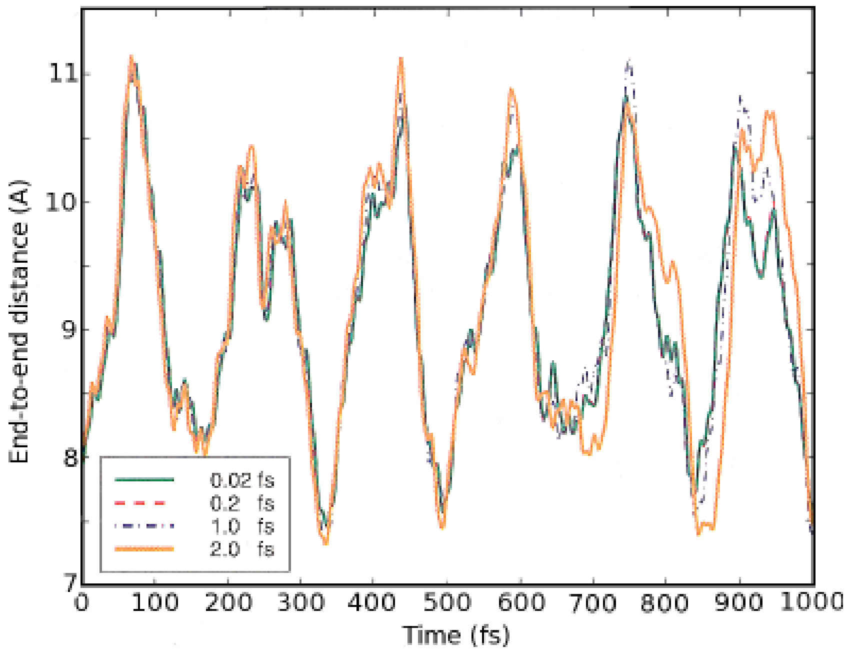
Max initial error for trajectory „vicinity“ in the time interval [0,tmax}, for the traj to stay within Δ_{max}.

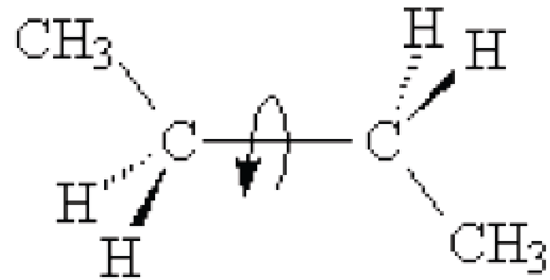
Example

$$e(t) = \left[\sum_i (r_i(t) - r_i^{ref}(t))^2 \right]^{1/2}$$



Trajectory vs. Timestep

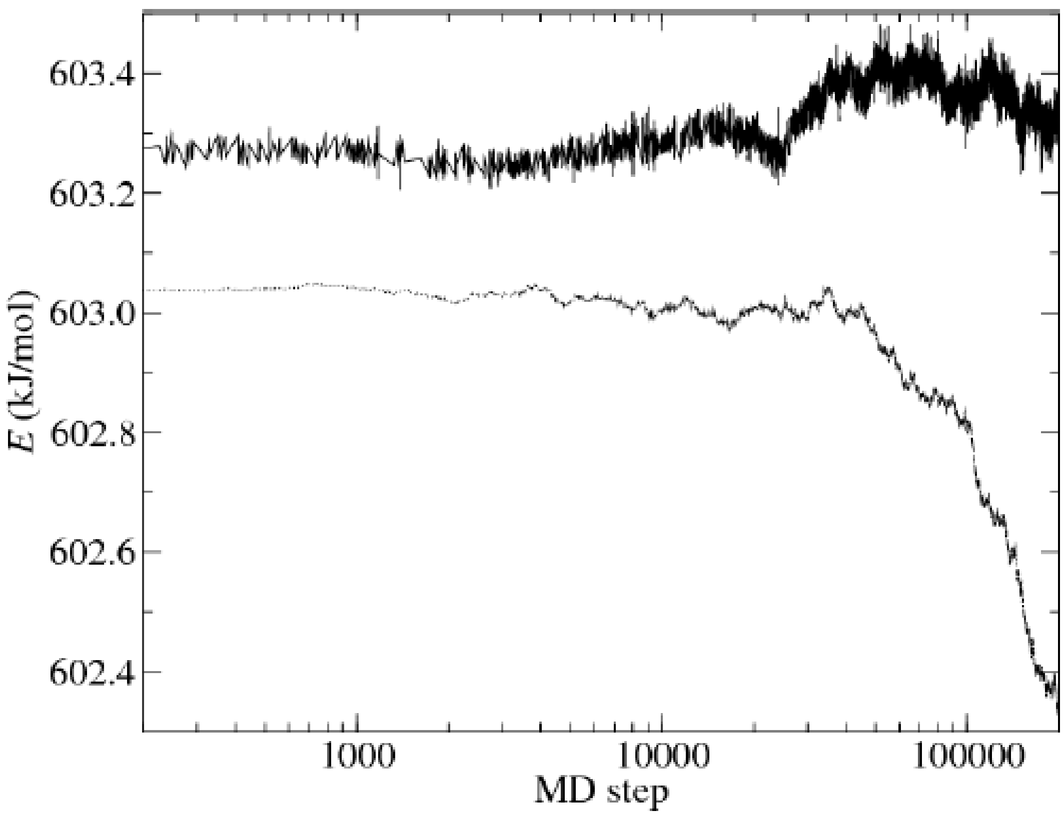




- After a time Δt, the trajectories become essentially uncorrelated.

Energy Conservation

Lennard-Jones liquid at 300K ($\Delta t = 1.5$ LJ units)



MD - Forces

classical MD simulations are based on a potential (force field), represented as a sum of pair potentials:

$$U(r^N) = \sum_{i<j=1}^{N} u(|r_i - r_j|)$$

where U(inf)=0.

Pairwise additivity, approximation. Such potentials are referred to as “non-bonding” potentials.

Very common is the Lennard-Jones potential: $4\varepsilon\left[\left(\frac{\sigma}{r^{12}} - \frac{\sigma}{r^{6}}\right)\right]$

attractive for $r < r_0$, repulsive for $r > r_0$.

The minimum represents “equilibrium” between rep. and attractive terms

The attractive term is asymptotic to r^{-6}.

MD-Potential: LJ

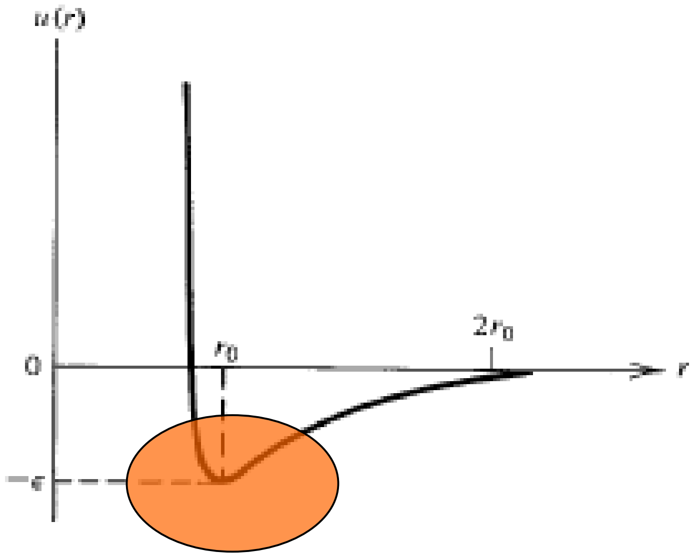
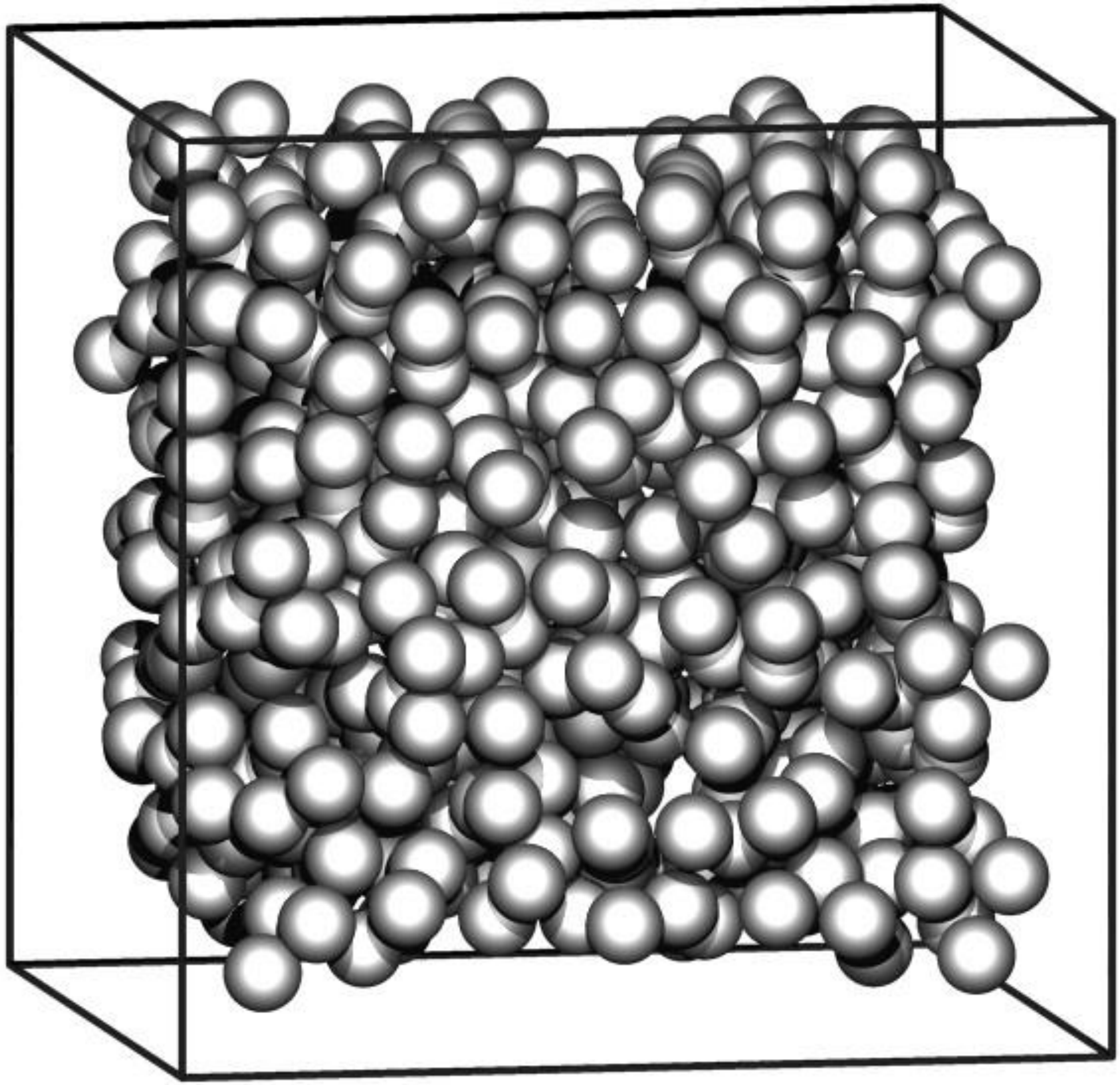


Fig. 7.6. A pair potential.

$$\langle E\rangle = \left\langle K(p^N)\right\rangle + \left\langle U(r^N)\right\rangle$$

$$= N\left\langle p^2/2m\right\rangle + \left\langle \sum_{i>j=1}^{N} u\left(\left|r_i - r_j\right|\right)\right\rangle$$

$$\langle E\rangle/N = \frac{3}{2}k_B T + \frac{1}{2}\rho\int d\boldsymbol{r}\; g(r)\, u(r)$$



http://matdl.org/repository/eserv/matdl:324/web_a_LJ_liquid.jpg

Quantities in MD

- Conserved quantities
  - T + V
  - Momentum (linear, angular)
- (Time) Averaged Quantities
  - Volume
  - Structure
  - Transport (diffusion, conduction)
- Sampled Quantities (see next lecture)
  - Free energy

Kinetic + potential energy

(vertically falling object)

→ constant

$$\frac{1}{2}mv^2 + mgh = konst.$$

Proof

$$\frac{dT}{dt} = \frac{d}{dt}(\frac{1}{2}mv^2) = \frac{1}{2}2mv\frac{dv}{dt} = mv\frac{dv}{dt} = Fv$$

$$\frac{dT}{dt} = Fv$$ (also as a vector!)

Since F=-mg:

$$F = -mg, v = \frac{dh}{dt}$$

Rate of change of kinetic energy is equal (and opposite in sign) to the rate of change of potential energy

$$\frac{dT}{dt} = -mg\frac{dh}{dt}$$

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

$$E_t = T + V$$

$$V = -\int F\,dx$$

$$\dot{V} = -\frac{d}{dt}\int F\,dx = -\frac{dx}{dt}\frac{d}{dx}\int F\,dx = -\dot{x}F$$

$$\dot{E}_t = \dot{T} + \dot{V}$$

$$\dot{T} = \frac{d}{dt}\left(\frac{m\dot{x}^2}{2}\right) = m\dot{x}\ddot{x}$$

$$E_t = T + V = const$$
$$\dot{E}_t = m\dot{x}\ddot{x} - \dot{x}F = 0$$

Forces - LJ

$$f_x(r) = -\frac{\partial u(r)}{\partial x} = -(\frac{x}{r})(\frac{\partial u(r)}{\partial r}).$$

$$u(r) = 4\left[\left(\frac{1}{r^{12}} - \frac{1}{r^6}\right)\right]$$

$$f_x(r) = \frac{48x}{r^2}(\frac{1}{r^{12}} - 0.5\frac{1}{r^6}).$$

$$F_x = -\partial U(r) / \partial x,$$

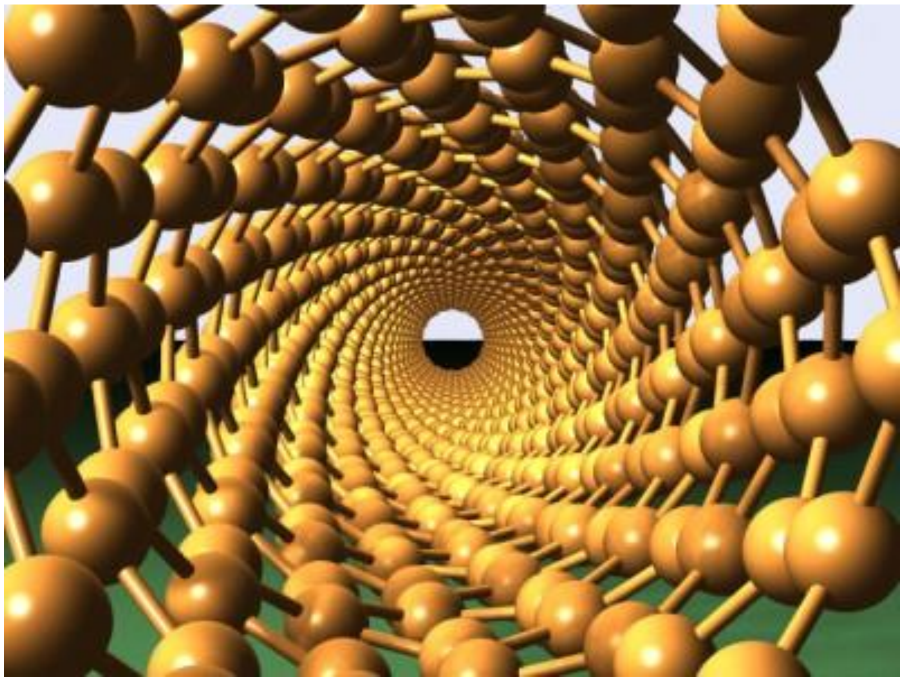
$$F_y = -\partial U(r) / \partial y,$$

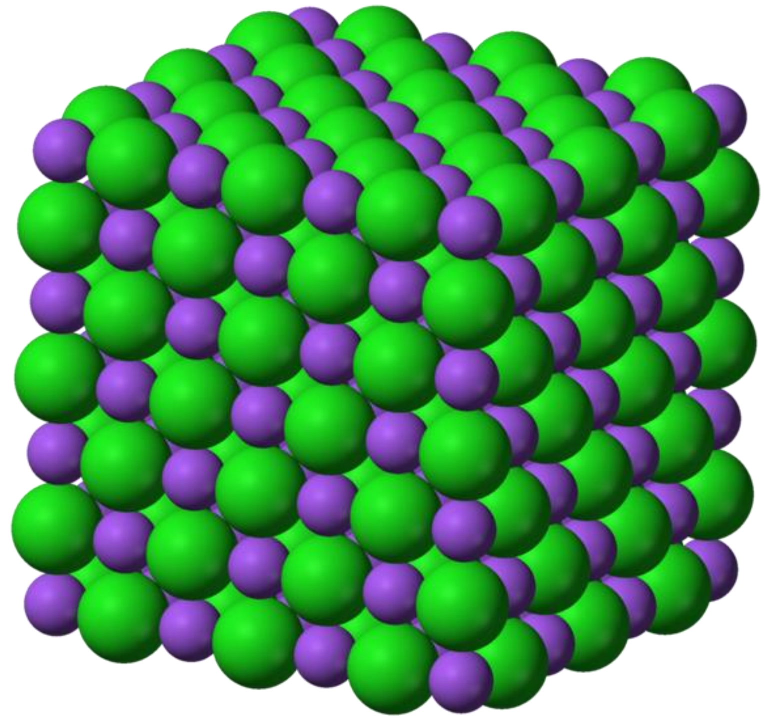
$$F_z = -\partial U(r) / \partial z$$

Forces - Algorithm

- Loop (i); 1, N-1
  - Loop(j); i+1, N
  - dist=x(i)-x(j)
  - (test periodic boundary conditions)
  - r2=dist^2
  - if (r2 < cutoff)
    - r2i= 1/r2
    - r6i=r2i^3
    - force=48*r2i*r6i*(r6i-0.5)
    - f(i)=f(i)+force*dist
    - f(j)=f(j)-force*dist
    - energy=energy+4*r6i*(r6i-1)-ecut
  - Loop j done
- Loop i done

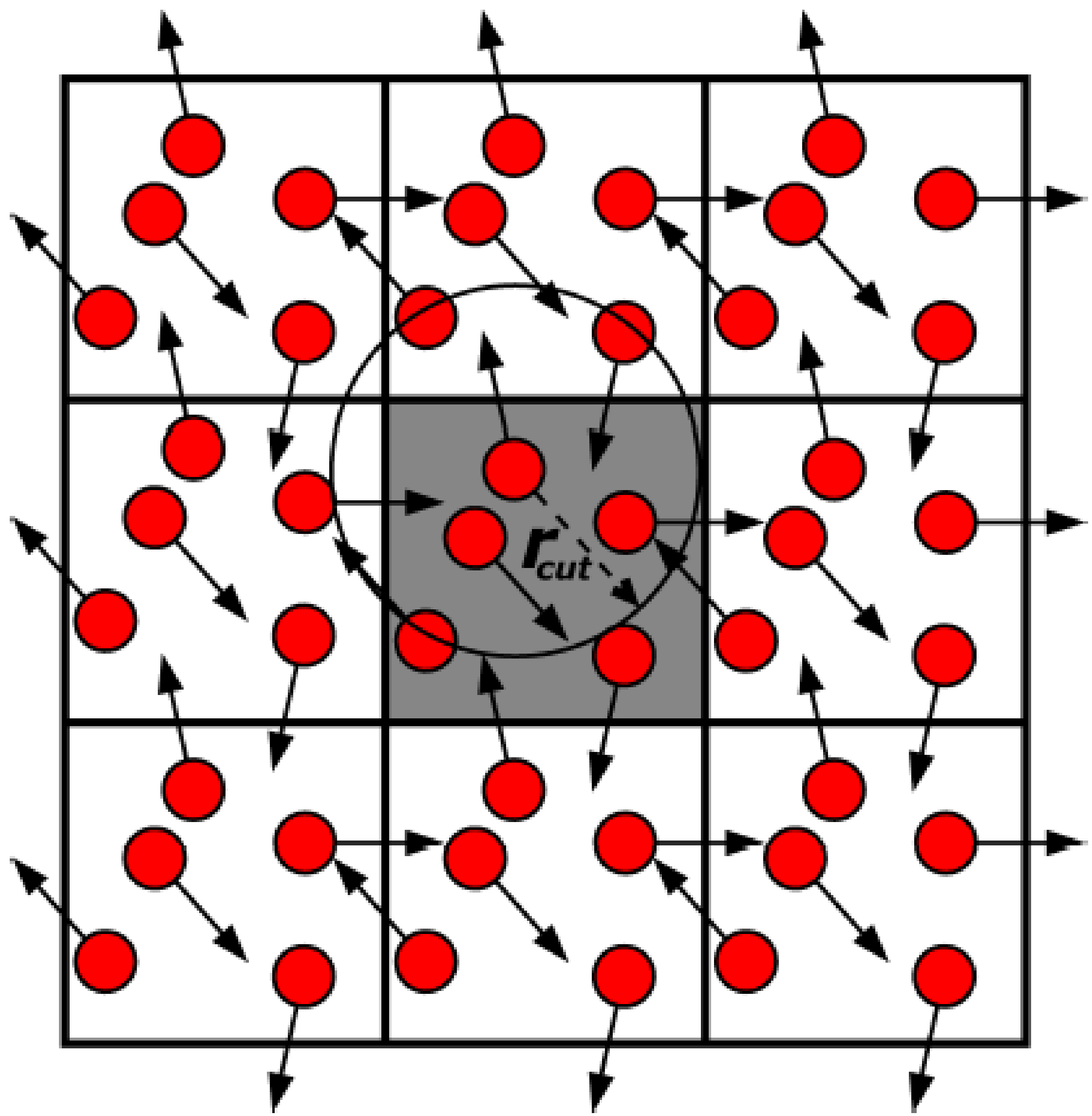
Infinite ?



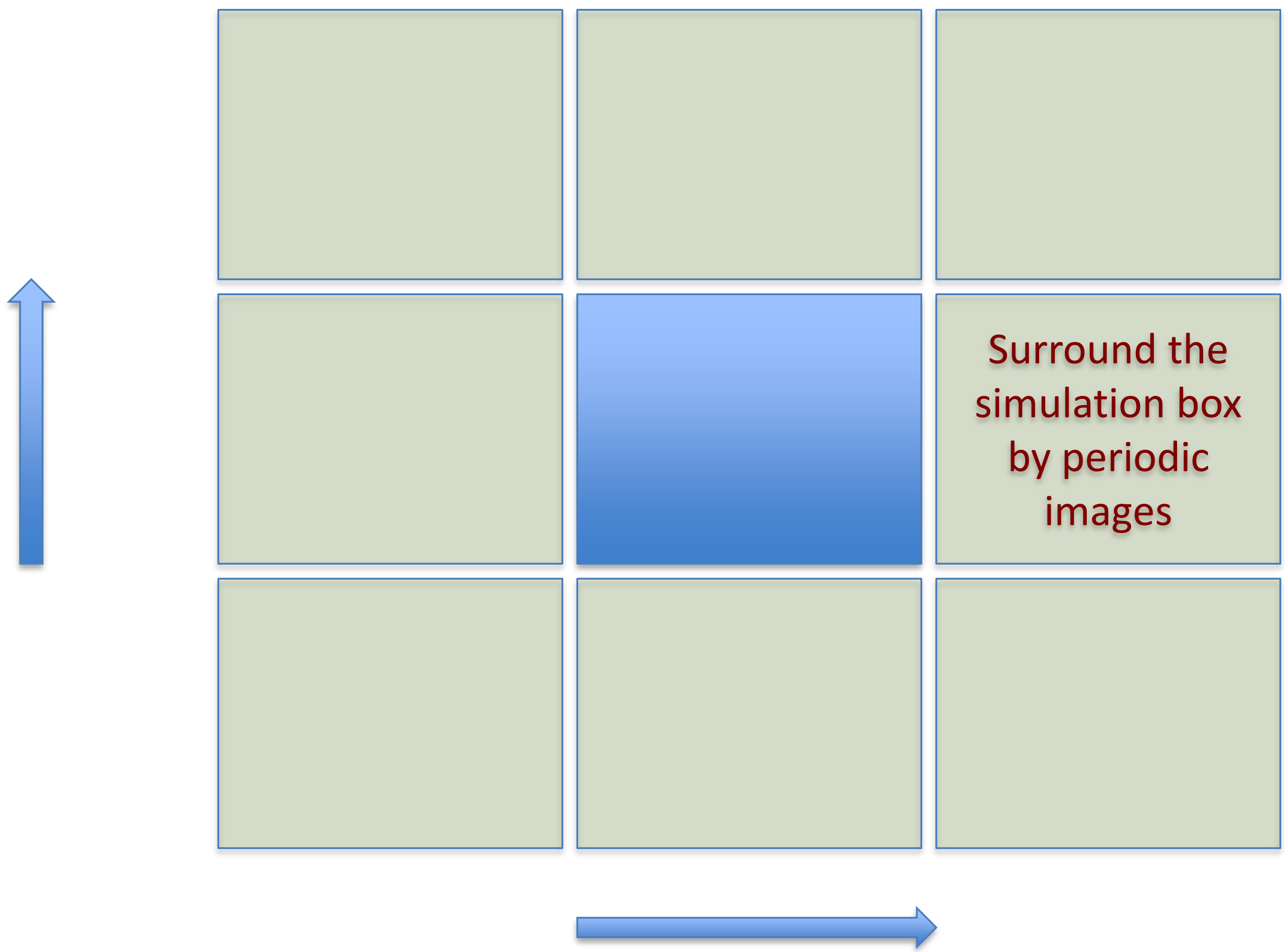


Molecule/Cluster

3D Crystal

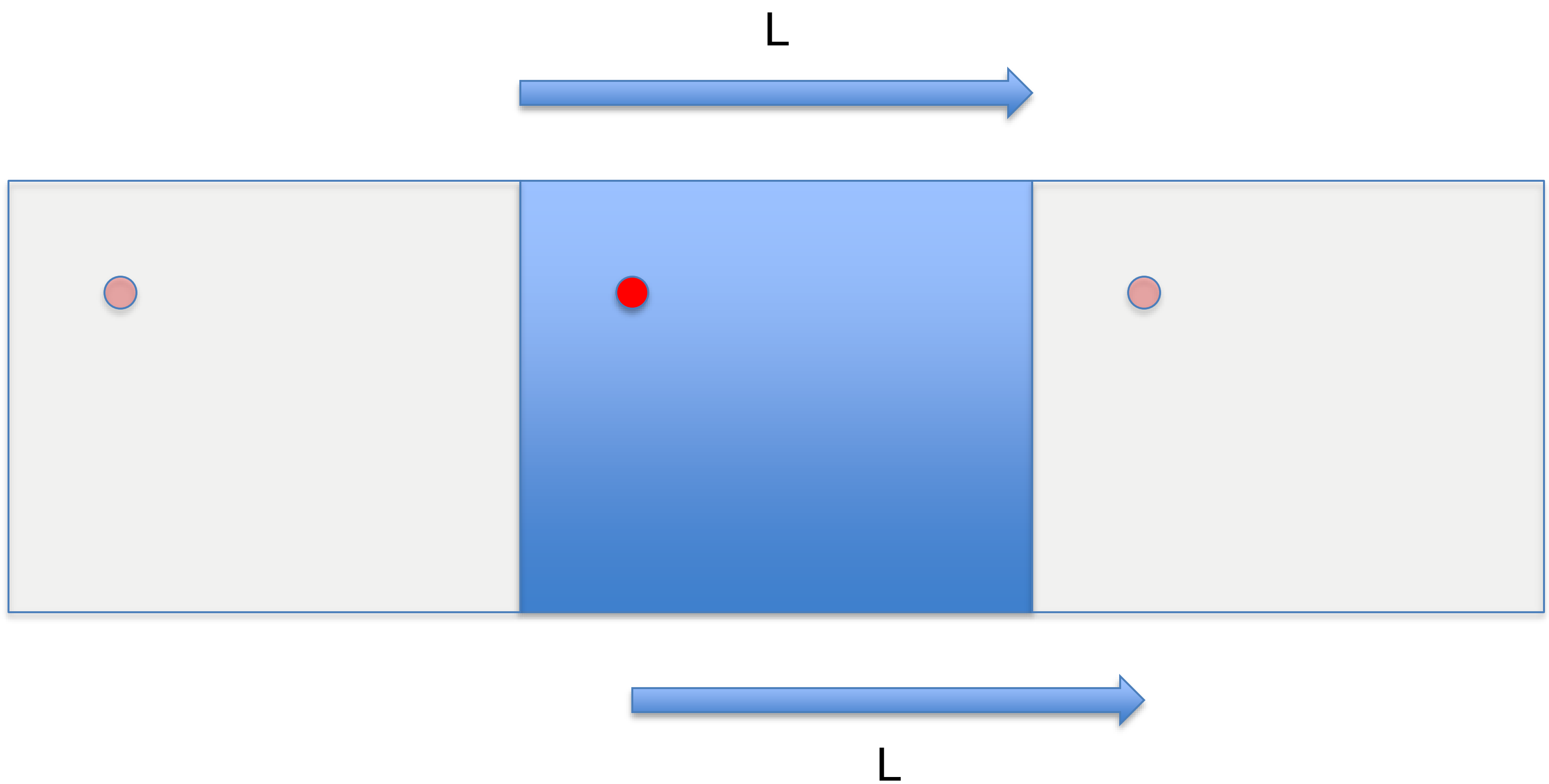

r_cut

Periodic Images



Simulation of “infinite” systems

periodic boundary conditions



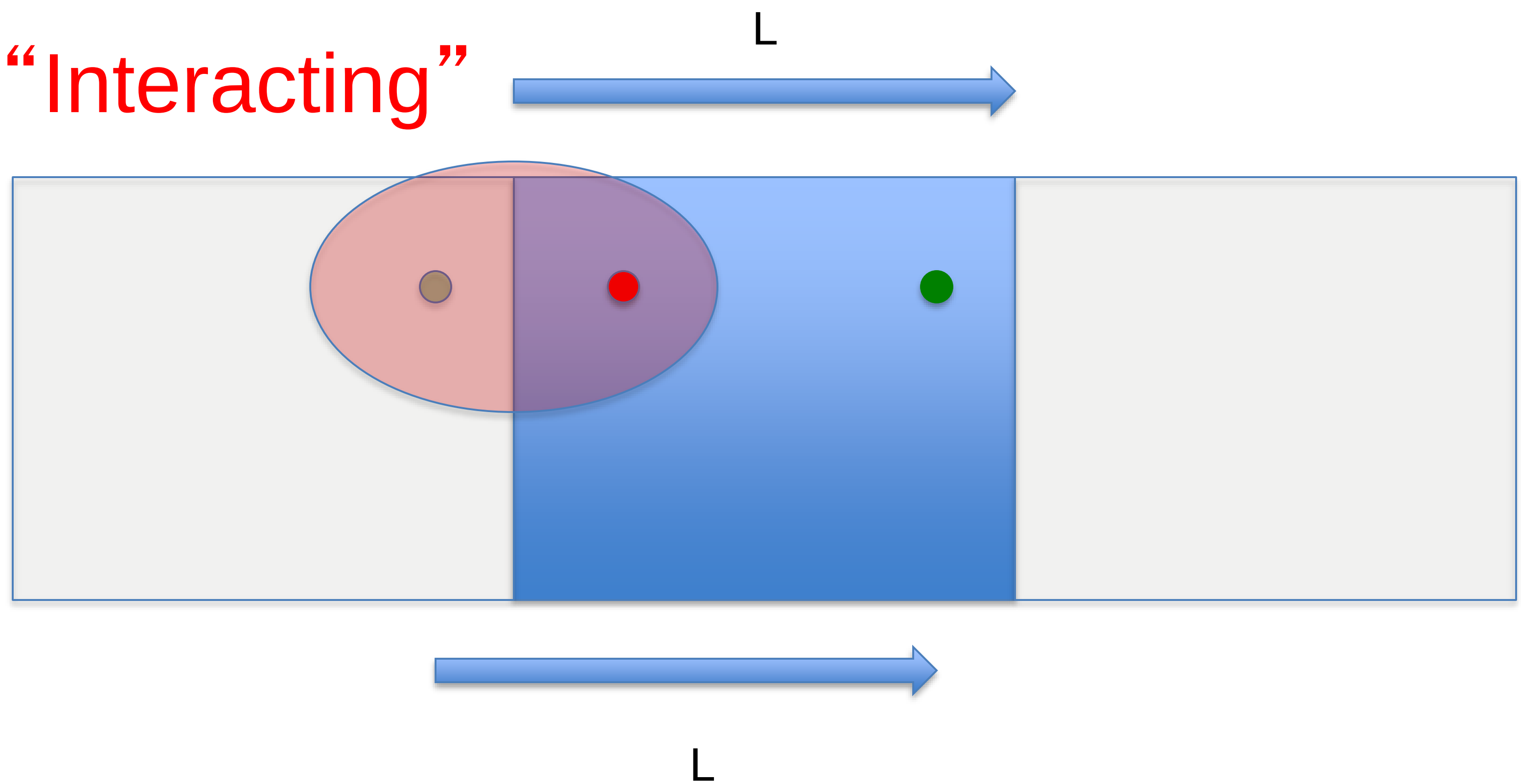
PBC

```
If (L-periodic) then

    if (x <   -L/2) x = x + L
    if (x >=  L/2) x = x – L

endif
```

Efficient realization:

$$x_i = \mathrm{mod}(x_i, L)$$

PBC & Distances



PBC & Distances

```
If (L-periodic) then
    dx = x(i) – x(j)
     if (dx >     L/2) dx = dx  - L
     if (dx <=  -L/2) dx = dx + L
endif
```

Efficient realization:

```
dx = x(i) – x(j)
dx = dx – nint(dx/L)*L
```

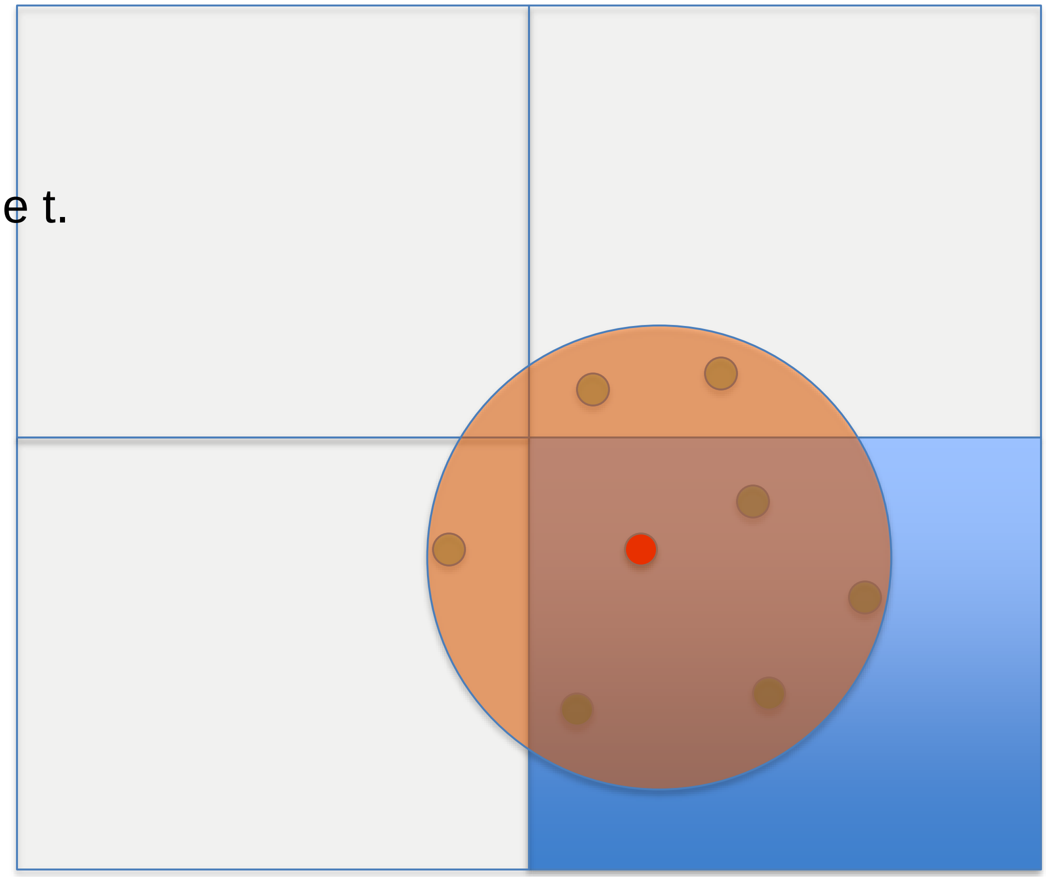
Verlet List – The Idea

Construct a list of adjacent atoms at time t.
Update the list at each step.

Advantages:

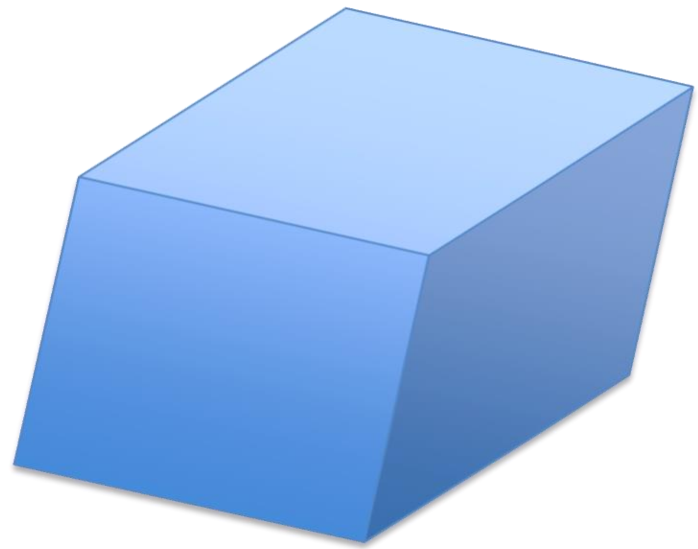
many distance vectors can be filtered out from the very beginning.

Shaded area: force field cutoff distance

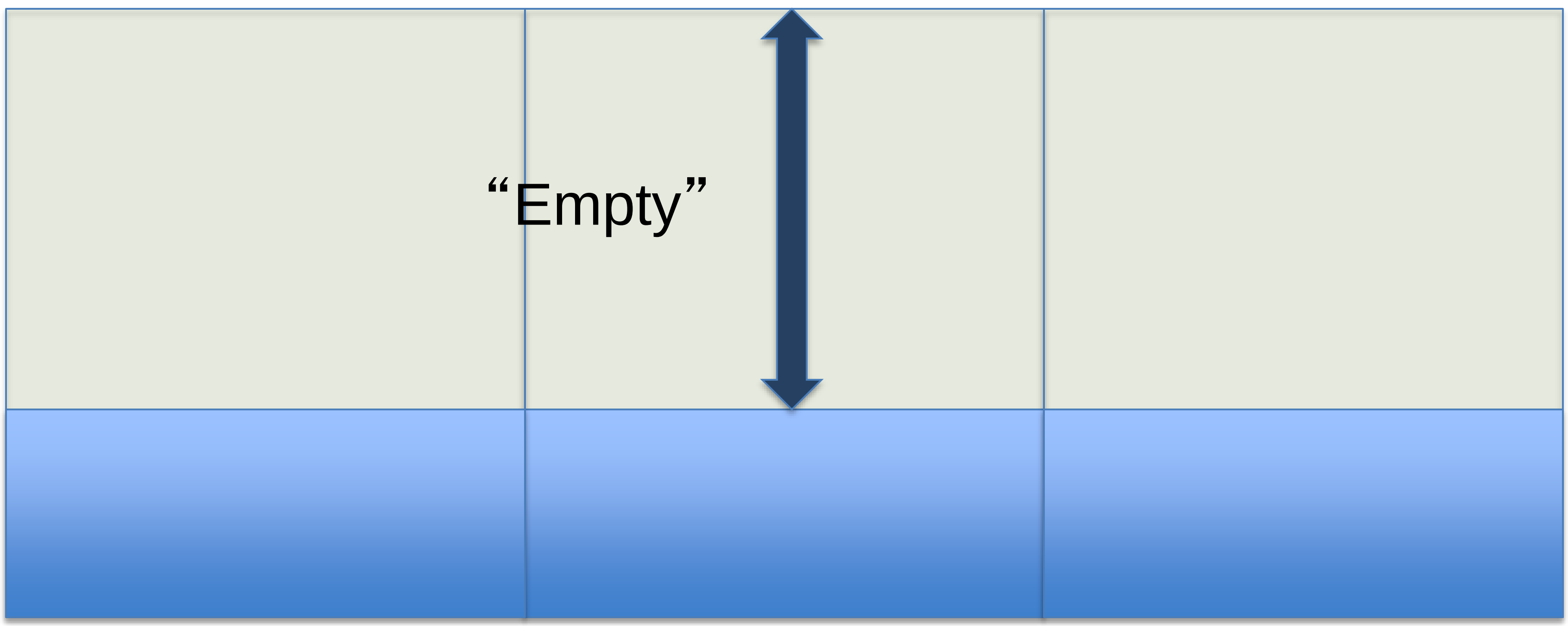


Geometries - Boxes

2D, 3D

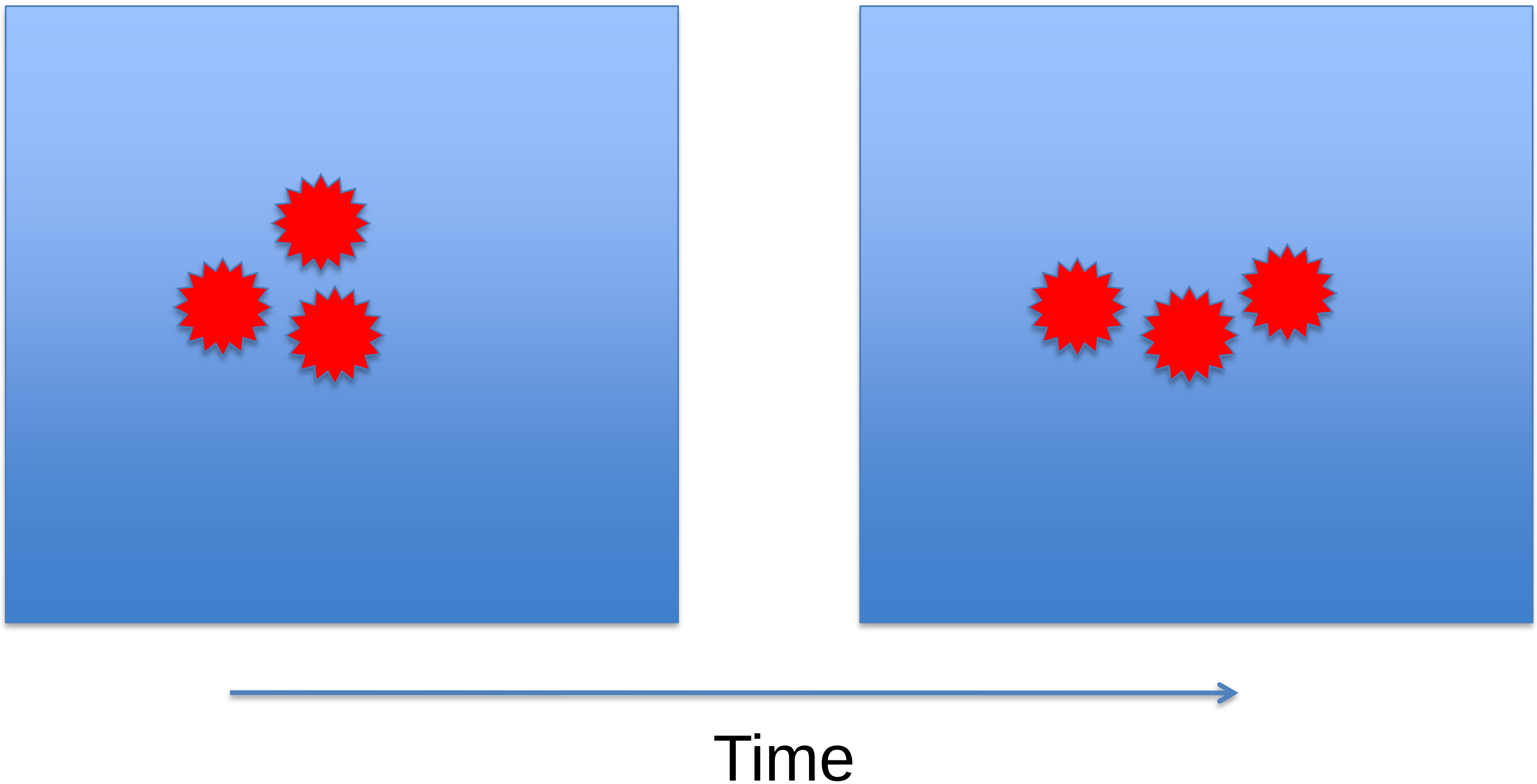


Slabs



Infinite surface

Molecules, Clusters



4

Charges, Force Field Models, More Properties, Machine Learning

SHORT-RANGE VS.
LONG-RANGE interactions

Truncation

- Short-range vs. long-range interactions: $r<r_{cut}$

$$U^{tot} = \sum_{i<j} u_c(r_{ij}) + \left\{ \frac{1}{2}\int_{r_c}^{\infty} dr \rho(r) u(r) 4\pi r^2 \right\}$$

Tail term

Valid for rapid decay of potential energy,
Implies g(r) ~ 1 for r > r_c.

What are te consequences of considering interactions within a given cutoff ?

Lennard-Jones - truncation

$$u^{tail} = \frac{1}{2} 4\pi\rho \int_{r_c}^{\infty} dr r^2 u(r)$$

$$= \frac{1}{2} 16\pi\rho\varepsilon \int_{r_c}^{\infty} dr r^2 \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$$

$$= \frac{8}{3} \pi\rho\varepsilon\sigma^3 \left[\frac{1}{3} \left(\frac{\sigma}{r_c} \right)^{9} - \left(\frac{\sigma}{r_c} \right)^{3} \right]$$

Define a tail correction – its value may be sizable due to large atom number at large r values!

General truncation strategies

- Hard cutoff

$$u^{truncated}(r) = \begin{cases} u^{lj}(r), r \le r_c \\ 0, r > r_c \end{cases}$$

- Shifted potential

$$u^{truncated-shifted}(r) = \begin{cases} u^{lj}(r) - u^{lj}(r_c), r \le r_c \\ 0, r > r_c \end{cases}$$

Ewald Summantion

- Charged particles (ions, partial charges)
- Long-range Coulomb interactions
- $O(N^{3/2})$, Ewald Summation
- Particle Mesh Ewald Summation (N logN) (1000-10.000 atoms)

- Point charges
- PBC

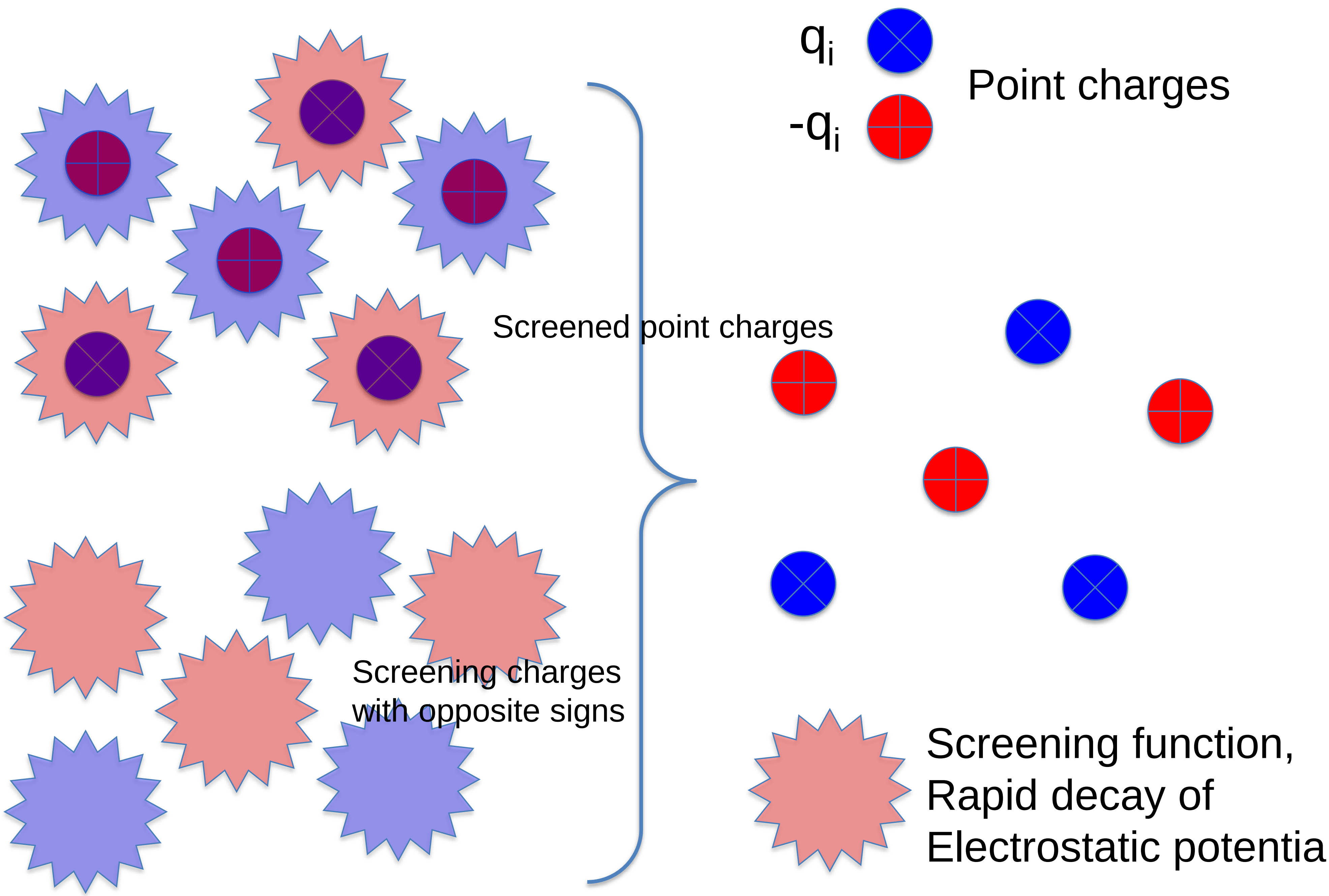
Ewald

Point charges

$$U_{Coulomb} = \frac{1}{2} \sum_{i=1}^{N} q_i f(r_i)$$

$$f(r_i) = \sum_{j,n}{}' \frac{q_j}{\left| r_{ij} + nL \right|}$$

Electrostatic potential at pos i


q_i
$-q_i$
Point charges
Screened point charges
Screening charges
with opposite signs
Screening function,
Rapid decay of
Electrostatic potential

Ewald, formula

Periodic sum of screening functions (Gaussians)

Reciprocal (Fourier) space

$$U_{Col} = \frac{1}{2V}\sum_{\mathbf{k}\neq 0}\frac{4\pi}{k^2}\left|\rho(\mathbf{k})\right|^2 \exp(-k^2/4\alpha) - (\alpha/\pi)^{1/2}\sum_{i=1}^{N} q_i^2 + \frac{1}{2}\sum_{i\neq j}^{N}\frac{q_i q_j \, erfc(\sqrt{\alpha}\, r_{ij})}{r_{ij}}$$

Constant correction term (depend on charges, not on Charge locations)

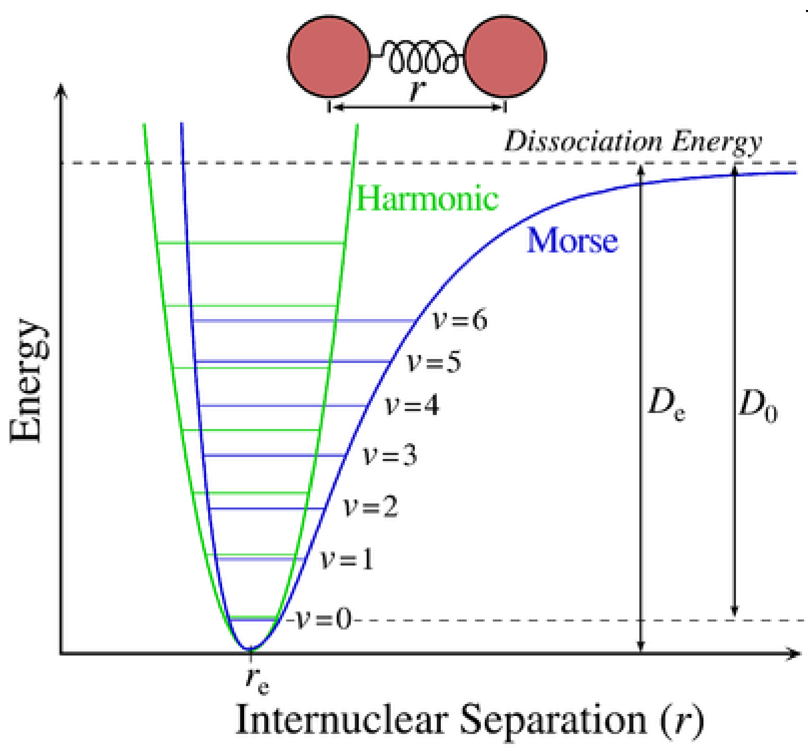
Electrostatic potential due to a charge q_i surrounded by a screening function (Gaussian) With net charge $-q_j$

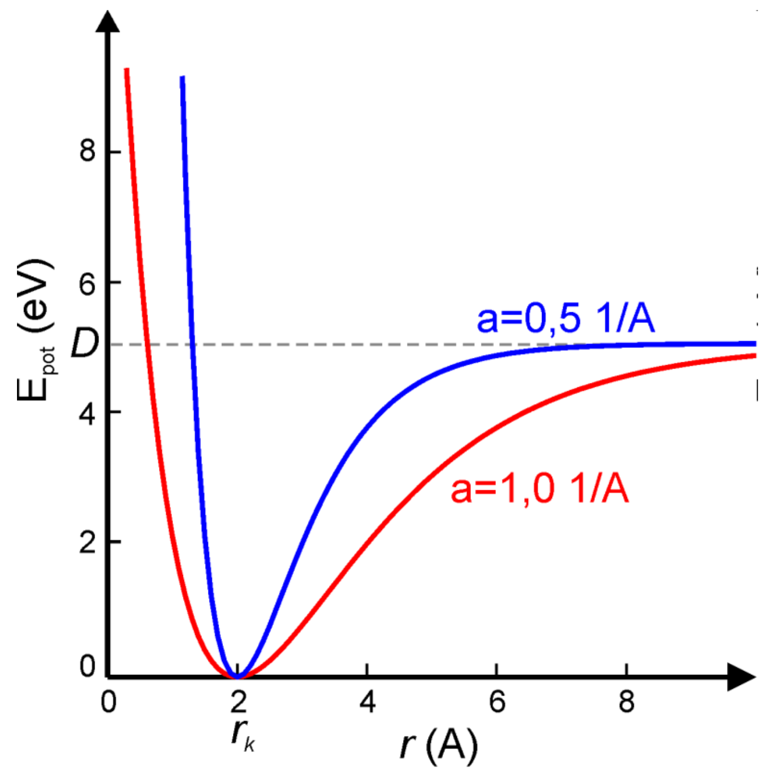
erfc(a) = 1 – erf(a)

*α, **k**=(k_x, k_y, k_z) are parameters*

Morse potential

$$U(r) = D_e(1 - e^{-a(r-r_e)})^2 - const$$





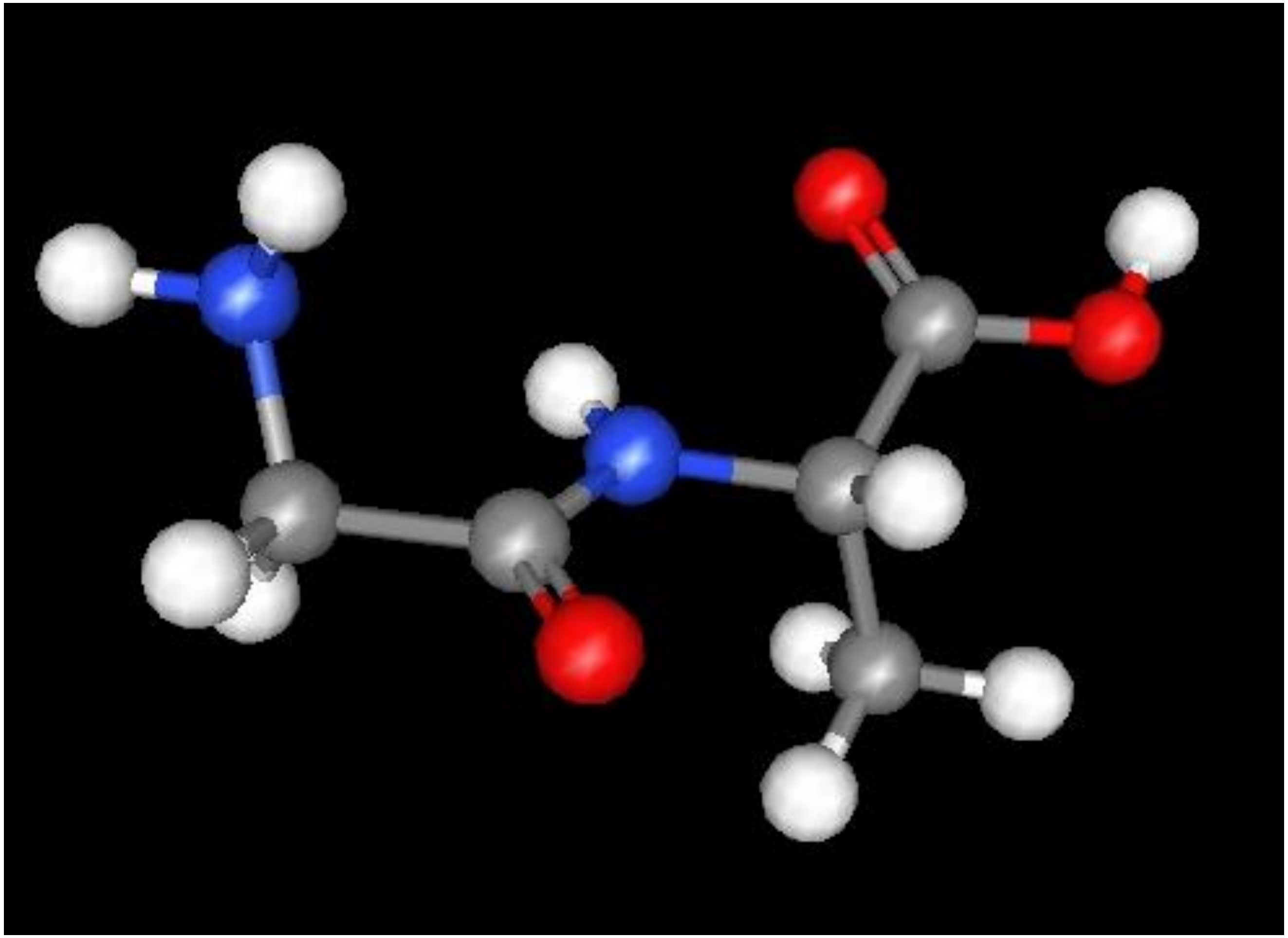
Amber Potential

$$U(r) = \sum_{bonds} \frac{1}{2} k_b (l - l_0)^2 + \sum_{angles} k_a (\theta - \theta_0)^2$$

$$+ \sum_{torsions} \frac{1}{2} V_n \left[1 + \cos(n\omega - \gamma)\right]$$

$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left\{ \varepsilon_{i,j} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$

Applications



Biological chemistry, proteins, macromolecules, also DNA

bhm (Born-Huggins-Meyer)

$$U(r_{ij}) = Ae^{\left[B\left((\sigma - r_{ij}\right)\right]} - \frac{C}{r_{ij}^{6}} - \frac{D}{r_{ij}^{8}}$$

Applications

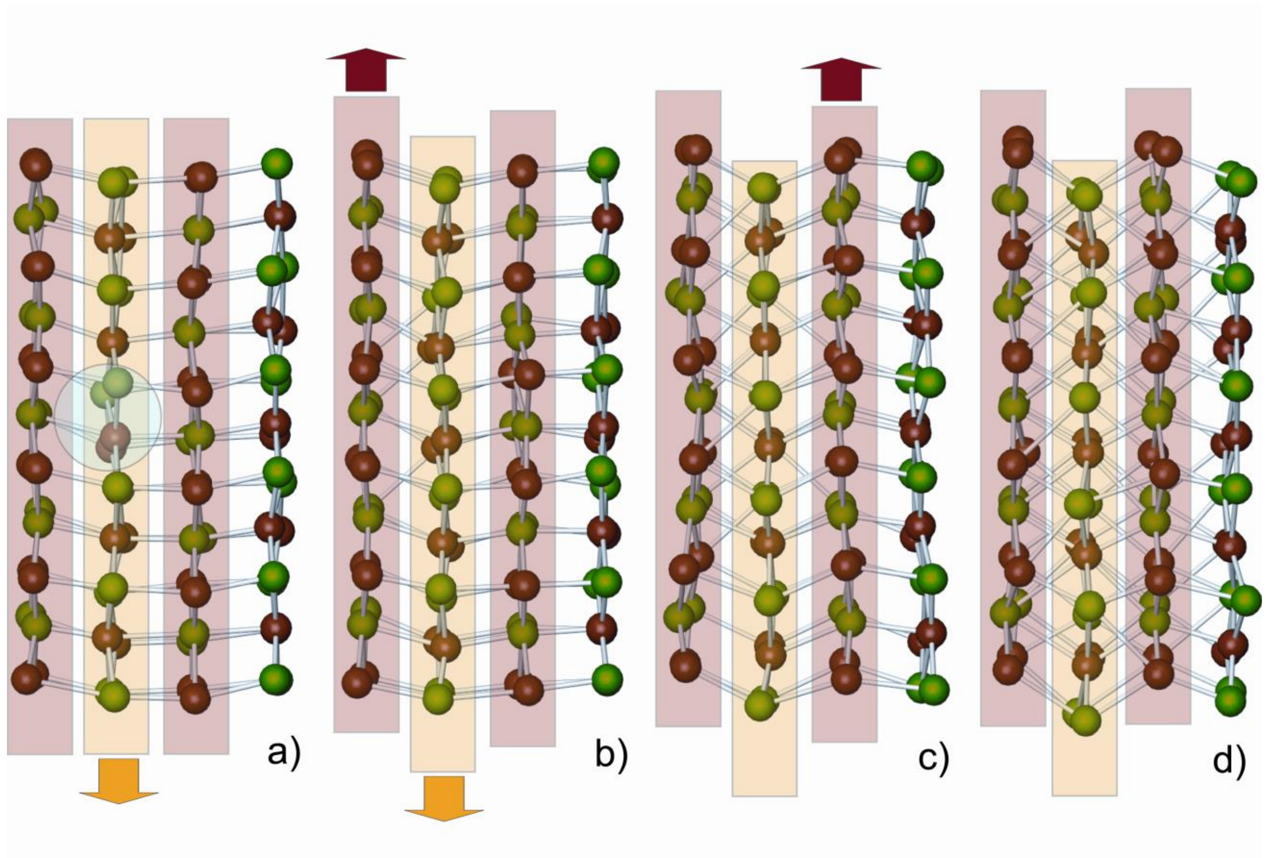
Ionic salts

- Bulk calculations
- Cluster calculations (also charged)
- Solvation, Condensation (solvent)

↓

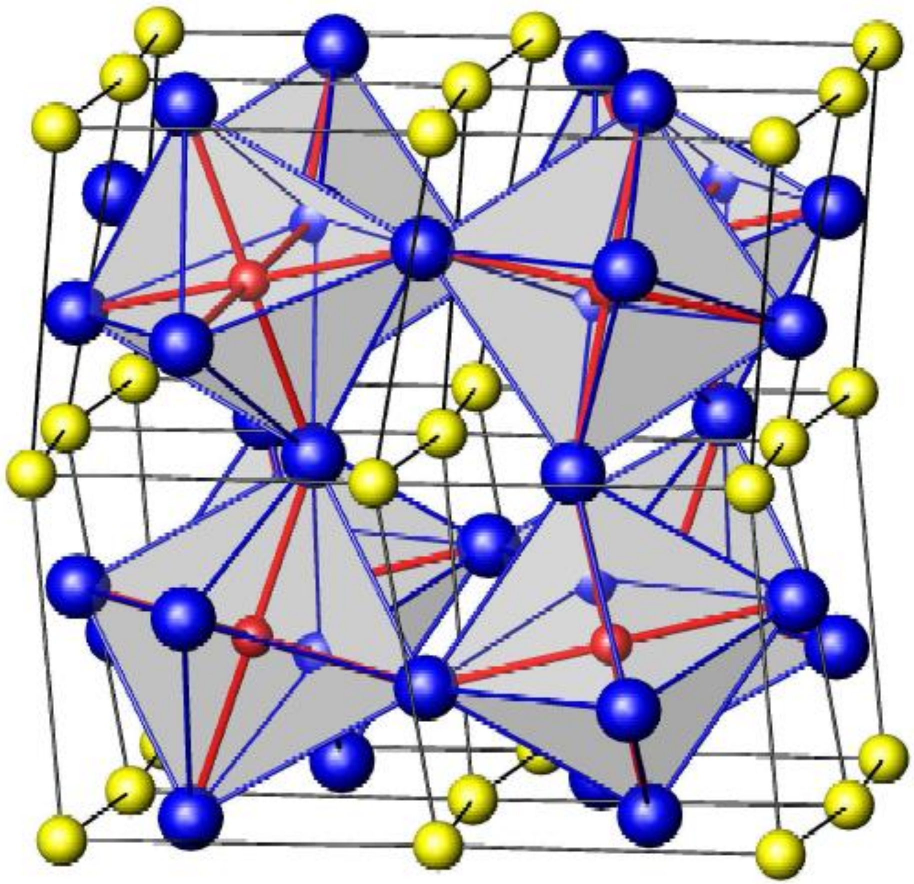
Fixed charges, typically integer numbers corresponding to oxidation numbers

Alkali halides



NaCl, KCl, ….

oxides



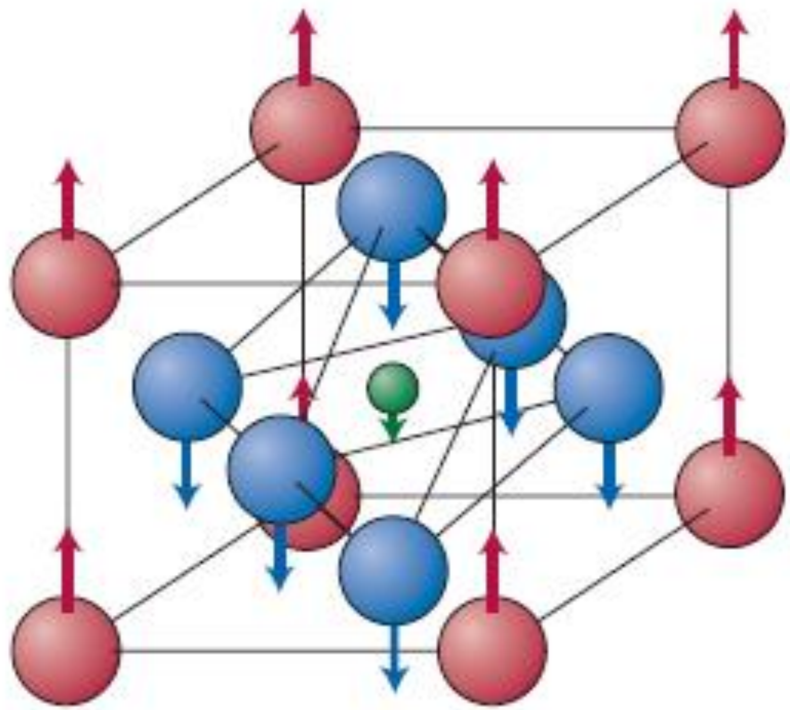
$CaTiO_3$
$BaTiO_3$

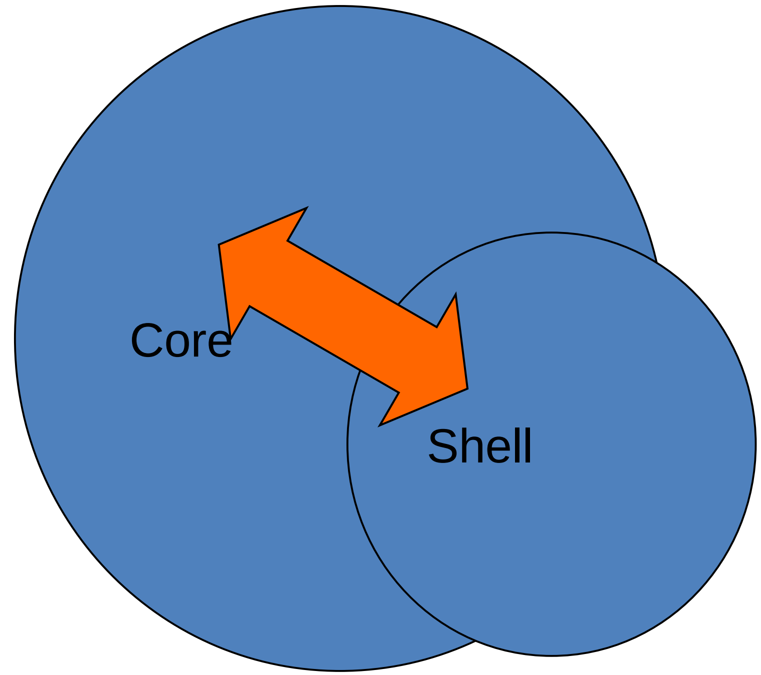
ZnO
GaN

…

Core-Shell Potential

An atom is represented as a combination of a core and a shell, both bearing a charge, and coupled through a harmonic potential (a spring).





Harmonic coupling

Polarization, anisotropic shapes
Example: Oxygen in Perovskites

Applications: batteries, fuel cells, ferroelectrics,...

Metal potential (EAM)

Many-body potential

$$U_{metal} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq 1} V_{ij}(r_{ij}) + \sum_{i=1}^{N} F(\rho_i)$$

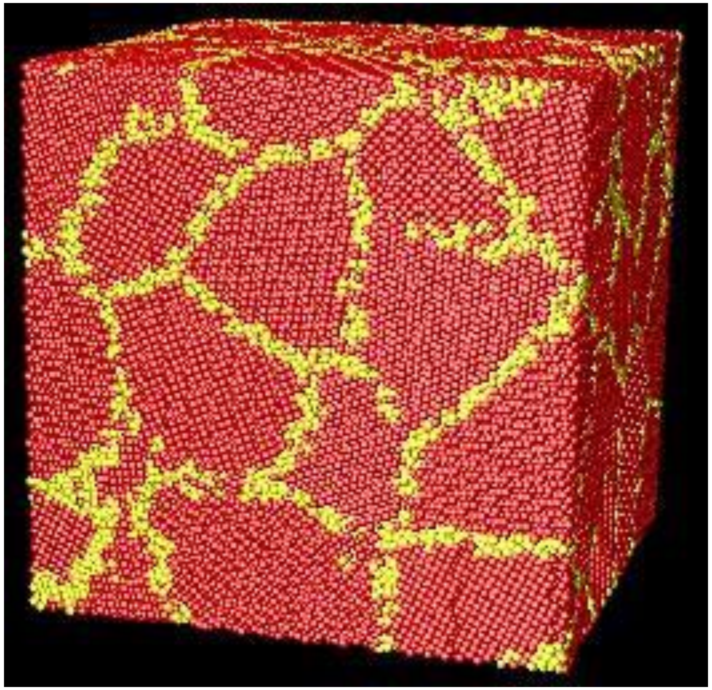
Two body term

Embedding Function

$$\rho_i = \sum_{j=1, j\neq i} \rho_{ij}(r_{ij})$$

Superposition of atomic functions
(EAM: spherical symmetric)

No charges, therefore no Ewald



Compound Classes:

Large (metallic) systems and alloys
Ni, Cu, Al, Fe, Pu
B, Ga
Ti, Zr

Bulk Structures and Surfaces

Typical Applications/Scenarios:

Nucleation
Domains
Mechanical properties
Shock-induced FT
Defects

Parameterization:

Empirical (based on properties)
DFT database

Finnis-Sinclair

$$V_{ij} = (r_{ij} - c)^2 (c_0 + c_1 r_{ij} + c_2 r_{ij}^2)$$

$$\rho_{ij} = (r_{ij} - d)^2 + \beta \frac{(r_{ij} - d)^3}{d}$$

$$F(\rho_i) = -A\sqrt{\rho_i}$$

Properties from Trajectories „computer experiments“

- Access experimentally measurable quantities from simulations;
- Average of (some function of) coordinates and momenta of the system particles;
- Directly accessible quantities: T, p, C_v
- Local properties: pair-distributuion g(r) ;
- Not directly accessible: free energy (*F*,*G*), *S*.

P, T, C_V

$$k_B T = \frac{\langle 2K \rangle}{f}, f = \text{ degrees of freedom}$$

$$p = \rho k_B T + \frac{1}{Vd}\left\langle \sum_{i<j} f(r_{ij}) \cdot r_{ij} \right\rangle$$

d is dimensionality of the system (2,3,…)

$$\left\langle K^2 \right\rangle_{NVE} - \left\langle K \right\rangle_{NVE}^2 = \frac{3k_B^2 T^2}{2N}\left(1 - \frac{3k_B}{2C_V}\right)$$

Equation of state

$$p = \rho k_B T - \frac{2}{3}\pi\rho^2 \int_0^\infty dr \frac{du(r)}{dr} r^3 g(r)$$

Diffusion

- D is a macroscopic quantity;
- It can be related to (stepwise) microscopic displacements;
- A "time integration" of the Δr can be a way of assessing diffusion in MD simulations.

r

$$\left\langle \mathrm{D}r(t)^2 \right\rangle = \frac{1}{N}\sum_{i=1}^{N} \mathrm{D}r_i(t)^2$$

Mean Squared Displacement

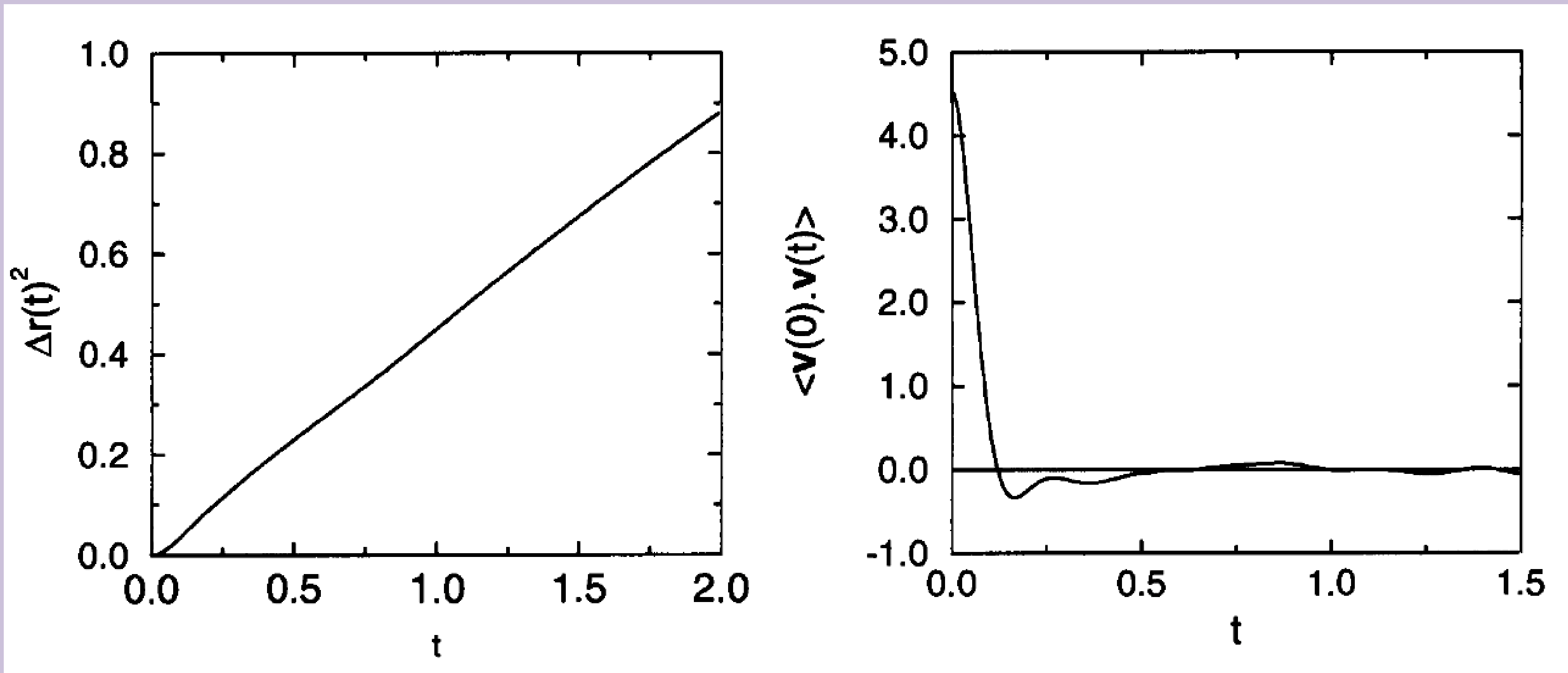
$$D = \frac{1}{6}\frac{MSD}{\mathrm{D}t}$$

p

$$D = \int_{o}^{\infty} dt \left\langle v_x(t) v_x(0) \right\rangle, t = \mathrm{D}t$$

Total elapsed time

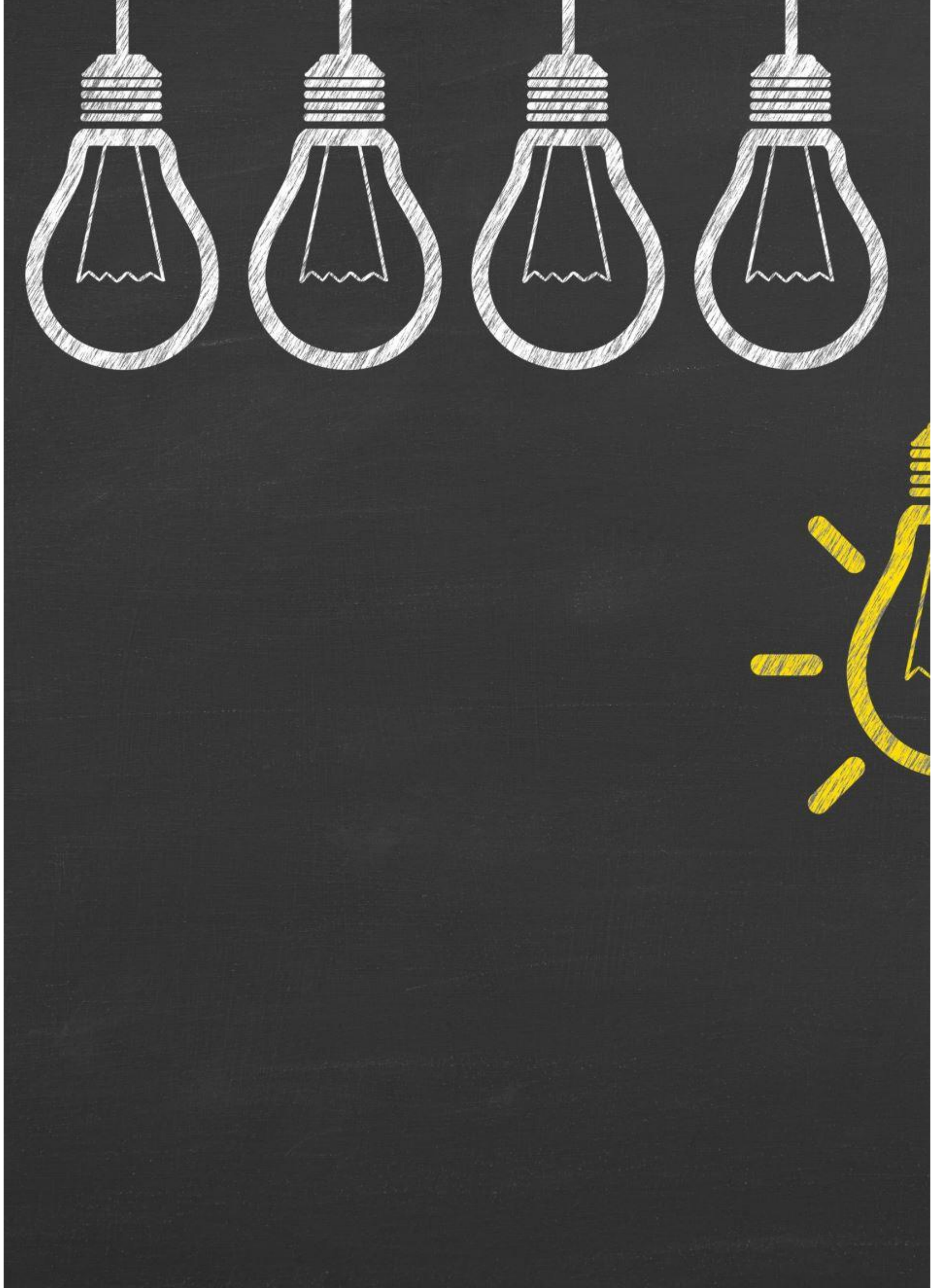
Velocity autocorrelation function



1.0
0.8
0.6
0.4
0.2
0.0
$\Delta r(t)^2$
0.0
0.5
1.0
1.5
2.0
t
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
<**v**(0).**v**(t)>
0.0
0.5
1.0
1.5
t

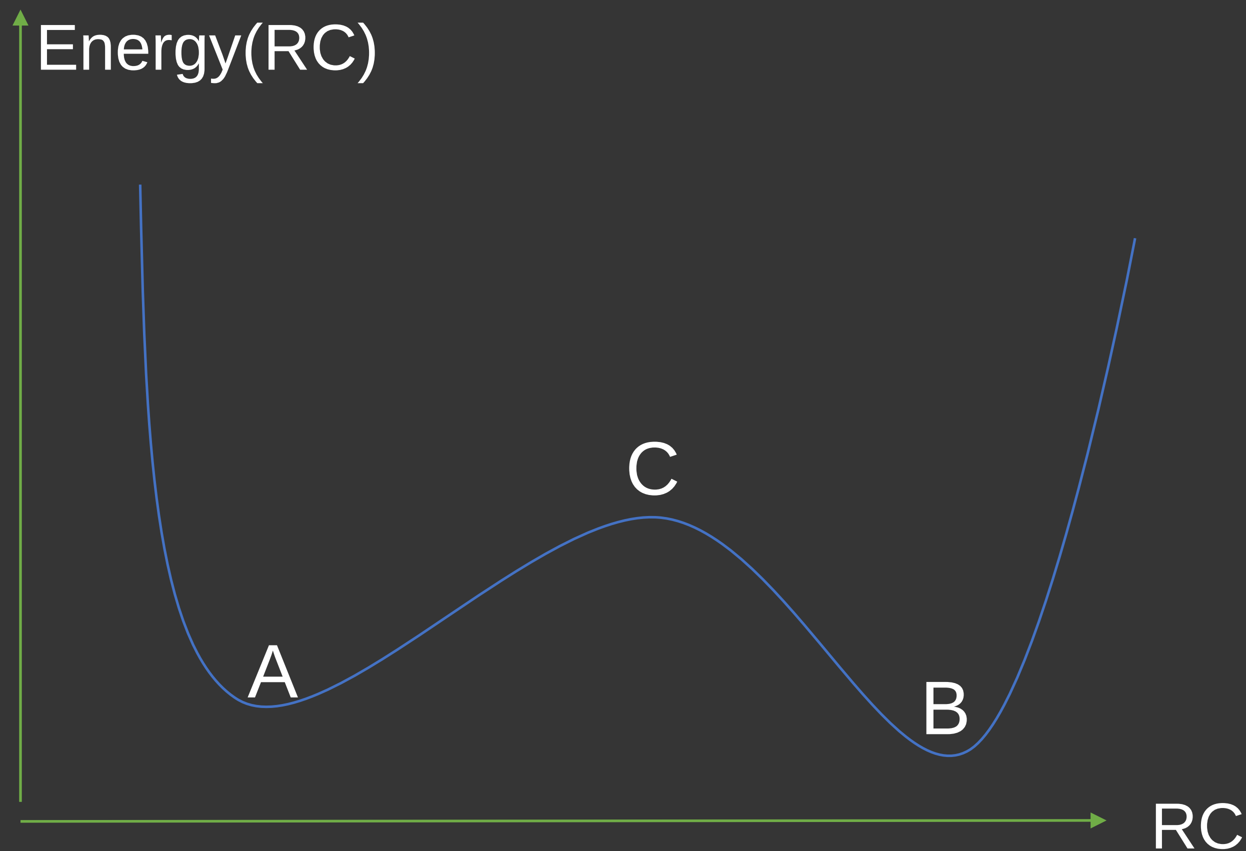
The empowering role of machine learning

Some ideas & why ML is here to stay



Potential Energy Surface

Statement of the Problem



Equation of Motion (nuclei)

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \int d\mathbf{r}\, \Psi^\star \mathcal{H}_\mathrm{e} \Psi$$

$$= -\nabla_I V_\mathrm{e}^\mathrm{E}(\{\mathbf{R}_I(t)\})$$

Learning Potential Energy Surfaces

Requirements

- High accuracy that is comparable to first principles methods, including high-order many-body effects,
- the ability to describe chemical reactions and arbitrary atomic configurations,
- simulation of large systems,
- a general applicability to all types of bonding and atomic interactions, from dispersion interactions via covalent bonds to metallic bonding,
- a strategy for systematic improvements, validation, and error control,
- a general automatic construction protocol,
- the absence of *ad hoc* approximations or system specific energy contributions which restrict the applicability to certain types of systems.

J. Behler, *Angew. Chem. Int. Ed*. **56**, 12828–12840 (2017).

Chartflow

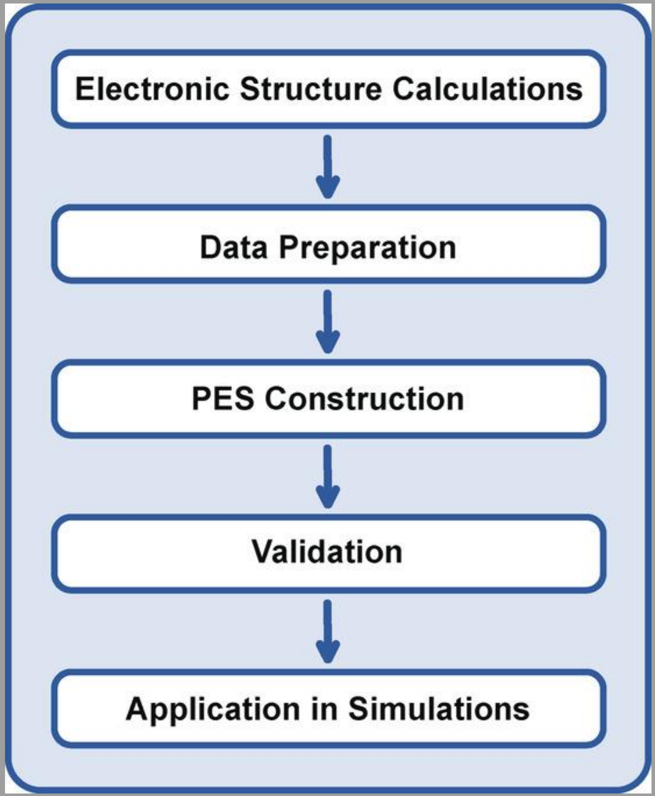
Calculation of the dataset

Preparation/standardisation

Learn

Test, then produce

In case, reopen and refit



J. Behler, *Angew. Chem. Int. Ed*. **56**, 12828–12840 (2017).

Neural Network Structure

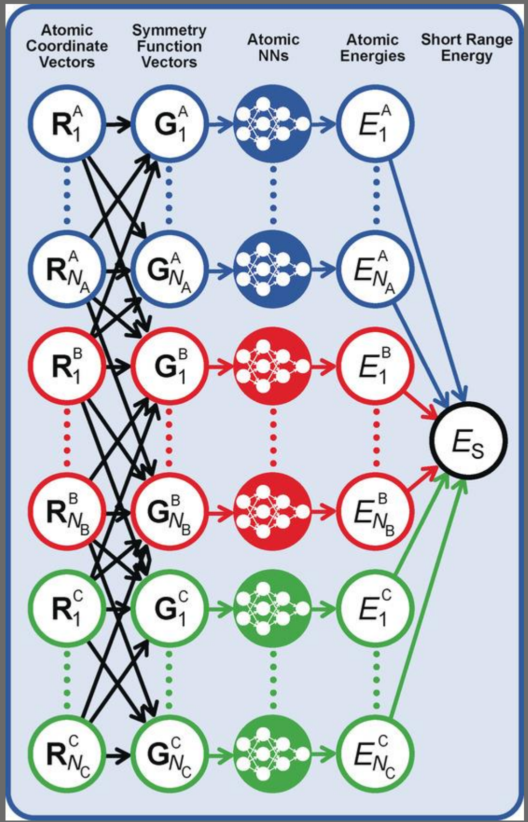
Coordinates;

Local chemical environment (symmetry functions);

Atomic Neural Networks;

Energy as sum of atomic energies;

Map of periodic structures by means of local atomic environments.



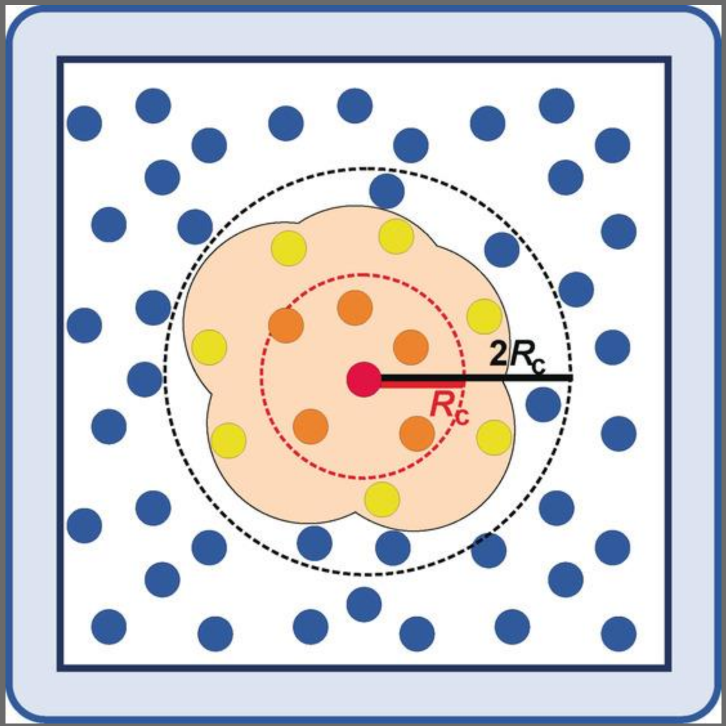
Atomic energy additivity, but higher precision than force fields.

Contribution to the force on an atom

Force on red atom depends on orange atoms & yellow atoms, since the latter are within the chemical environment of orange atoms;

Radius to be considered is $2R_c$;

Simplifications possible as contribution of yellow atoms may be small.



J. Behler, *Angew. Chem. Int. Ed*. **56**, 12828–12840 (2017).

Advantages for MD

- MD relies on the computation of forces;
- High precision enforces smaller systems and shorter simulation times;
- DFT ensures portability;
- Fore fields are specific and do not cover many-body terms, also they may fail if configurations are unseen;
- ML allows for high precision for large systems!

Study of Phase Diagrams - Silicon

Metadynamics Simulations of the High-Pressure Phases of Silicon Employing a High-Dimensional Neural Network Potential

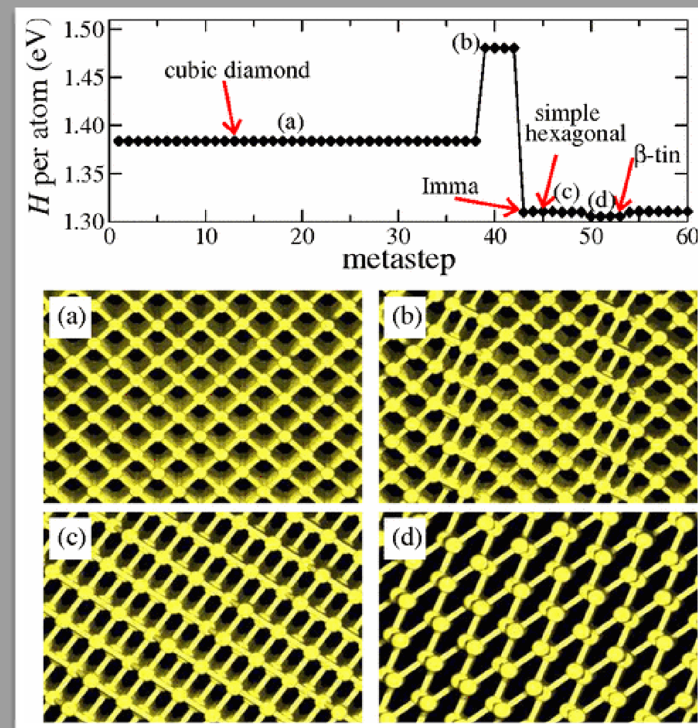
Jörg Behler, Roman Martoňák, Davide Donadio, and Michele Parrinello

Phys. Rev. Lett. **100**, 185501 – Published 5 May 2008

It needs a more efficient potential;

NN offers the best compromise between accuracy and performance (speed);

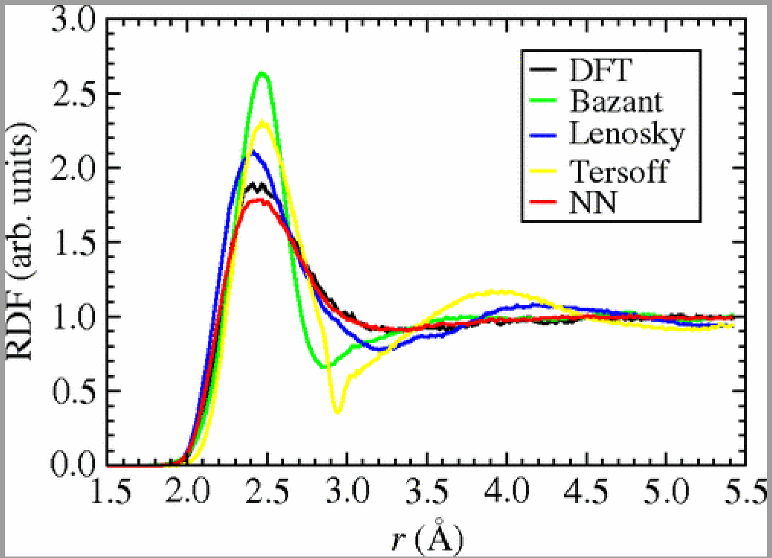
This has the advantage that phase diagrams can be systematically explored.



Quality of the fitting

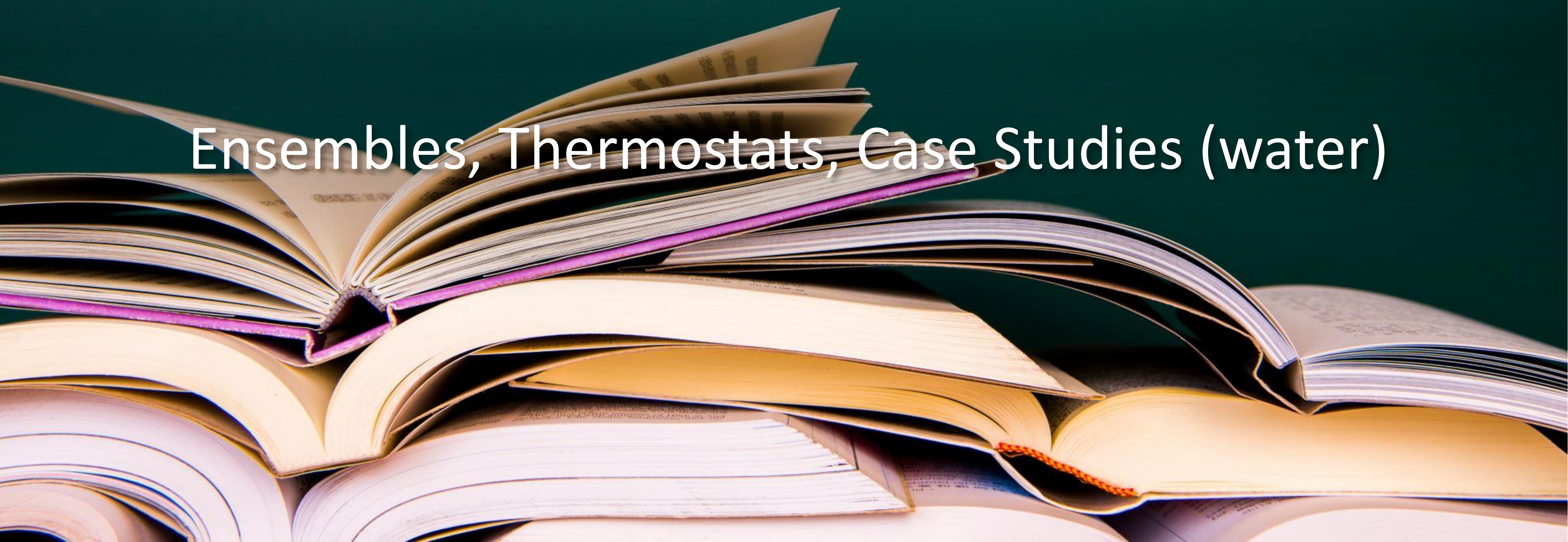
NN potential dramatically improves on the quality of the fitting to DFT data.

Simpler potentials (tersoff) may miss on important features, and therefore introduce errors.



5

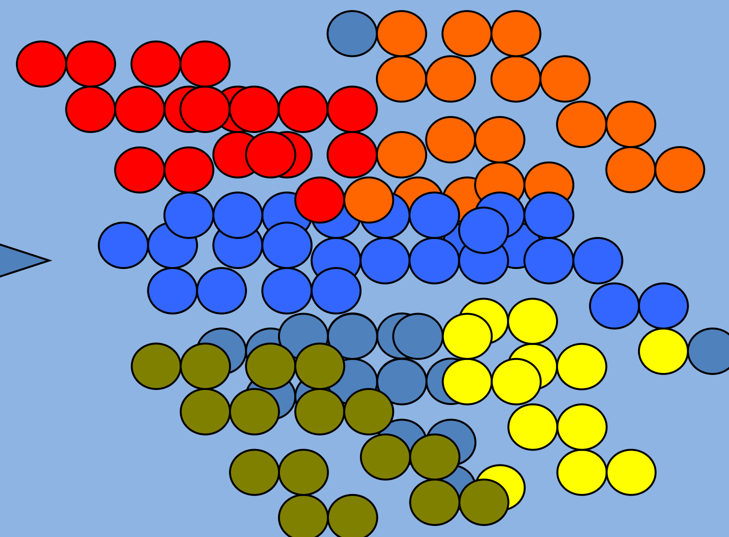
Ensembles, Thermostats, Case Studies (water)



- We need to establish a link between the time-evolution of a system and measurable macroscopic properties of a (many-body) system.

Statistical Mechanics

Atoms, molecules
Interparticle interactions;
Microscopic laws, QM.



Ensemble (many particles)
Law(s) of the ensemble ?

What are the observable properties of a system, if their interparticle interaction is governed by microscopic laws (i.e. quantum mechanics?)

Statistical Mechanics

The microscopic state of a many-body quantal system is determined by the Schrödiger equation (SE):

$$\mathcal{H}\,|\psi\rangle = i\hbar\,\frac{\partial|\psi\rangle}{\partial t}$$

With static solution:

$$\mathcal{H}\,|\psi_v\rangle = E_v|\psi_v\rangle$$

The index *v* is the collection of quantum numbers, given by *N* multiplied by the dimensionality, *D : v* = N x *D .*

Integration of the SE then provides the time evolution of the system, once the initial state is specified (see also Lecture 1)

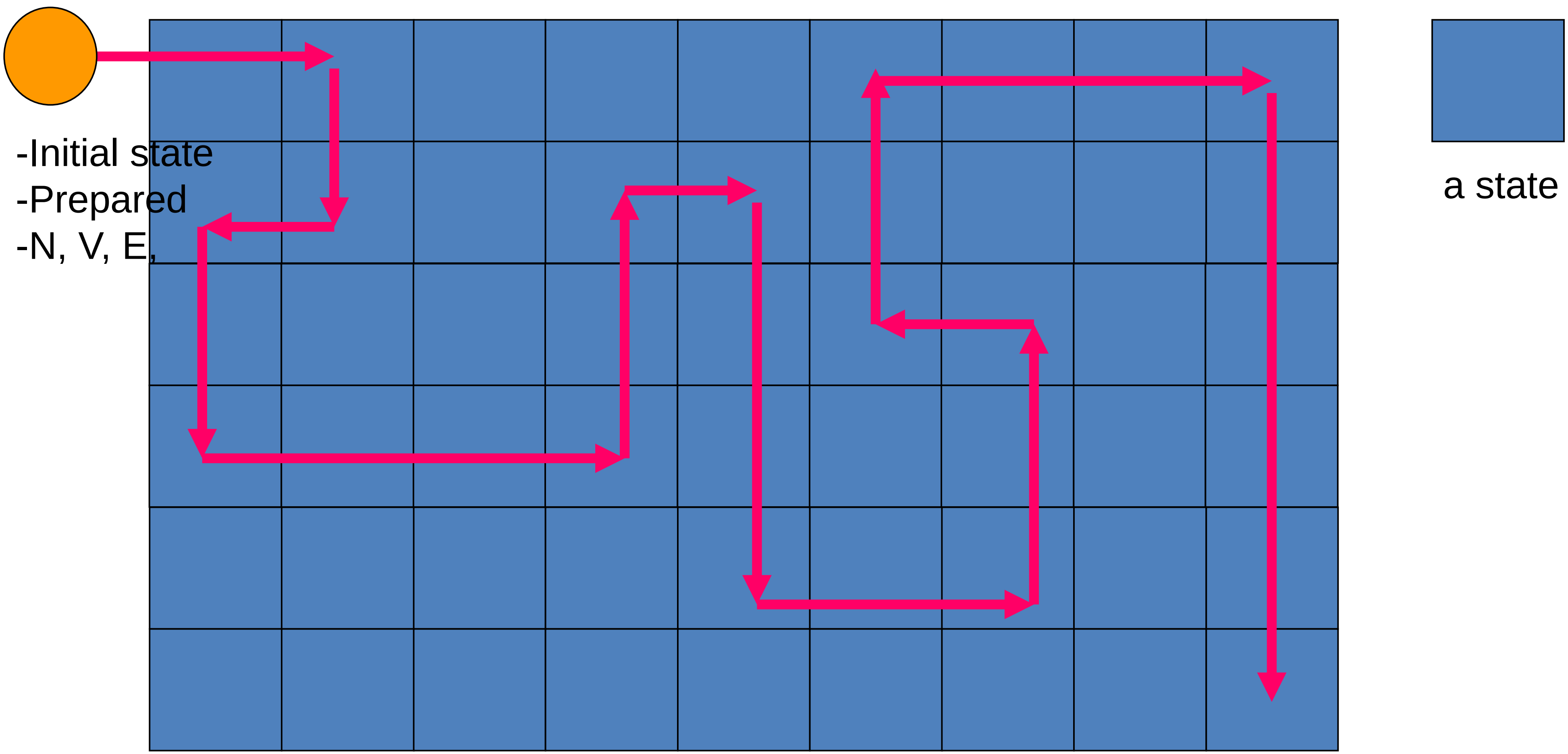
Classically, the mechanical state of a system is fully characterized by specifying points in phase space:

$$(r^N, p^N) = (r_1 \ldots r_N; p_1 \ldots p_N)$$

Flow in this space is determined by time integrating of Newton's equation of motion,

F = *m x a*, from an initial phase space point → a trajectory (see Lectures 1, 2).

Trajectory in *state* space (phase state or Hilbert space)



System preparation

- In preparing the system, a certain number of variables is chosen, like the number of particles *N*, the volume of the simulation box *V*, and the total energy of the system, *E*.
- This choice will define a surface in phase space, allowing to somehow reduce the large number of states that will be visited.
- This variable choice is summarised in the acronym *NVE*.

Trajectories in NVE phase space

- After some (long enough) time, the system will eventually visit all the microscopic states consistent with the constraints we have introduced to control the system. The final "measurement" will consist of a series of measurements on the system:

$$G_{obs} = \frac{1}{N}\sum_{a=1}^{N} G_a$$

G_a corresponds to the a^{th} measurement,
performed during a vanishingly short period of time, such that the system can be in only one microscopic state.

$$G_{obs} = \sum_{\nu} \left[\frac{1}{N} (\text{times state is in } \nu) \right] G_{\nu}$$

Probability or *weight* of state *v*

$$G_{obs} = \sum_{\nu} \frac{N_{\nu}}{N} \langle \nu | G | \nu \rangle$$

$$G_{obs} = \sum_{\nu} P_{\nu} G_{\nu} = \langle G \rangle$$

Ensemble average

Ensemble: assembly of all possible microstates, consistent with macroscopic constraints.
Microcanonical, NVE, assembly of all states with fixed total energy E and size N.
Canonical Ensemble, NVT, energy E can fluctuate (use of a thermostat, see below).

ergodicity

- Taken over a long period of time, the *ensemble average* and the *time average* are the same.
- Dynamical systems that obey this property are said to be *ergodic*.
- A system visits all possible states in time.
- Subdivision into subsystems, which are larger than the correlation length (subsystem uncorrelated): one instantaneous measure of the total macroscopic system is equivalent to many independent measurements of the macroscopic subsystems.

Microcanonical ensemble & the foundation of thermodynamics

The basic idea: every possible microscopic state or fluctuation does in fact occur, and observed properties are in fact the averages from all the microscopic states.

(i.e. we can measure equilibrium properties)

Assumption about the behaviour of a many-body system:

For an isolated system with total energy E, and given size (V, N), all microscopic states are equally likely at thermodynamic equilibrium.

$\Omega(N, V, E)$ = number of microscopic states with N and V, and E between E and E+δE.

(continuum approximation, E levels closely spaced)

$$P_v = \frac{1}{\Omega(V, N, E)}$$

For states in the ensemble, outside the ensemble P_v=0

Microcanonical Ensemble

$$S = k_B \ln \Omega(V, N, E)$$

Definition of Entropy

$$\frac{1}{T} = (\partial S / \partial E)_{N,V}$$

$$\beta = (k_B T)^{-1} = (\frac{\partial ln\Omega}{\partial E})_{N,V}$$

Temperature is positive, therefore Ω(N,V,E) is a monotonic Increasing function of E.

Constant Temperature MD simulations

$$W(p) = \left(\frac{b}{2pm} \right)^{3/2} \exp\left[-bp^2 / (2m) \right]$$

$$k_B T = m \left\langle v_a^2 \right\rangle$$

Maxwell-Boltzmann (M-B) velocity distribution governs the probability *W* of a momentum *p*.

Temperature of a particle.
Constant T is not equal to constant E_{kin}/particle!

Constant Temperature

$$\frac{s_{T_K}^2}{\langle T_K \rangle_{NVT}} = \frac{1}{N} \frac{\langle p^4 \rangle - \langle p^2 \rangle^2}{\langle p^2 \rangle^2} = \frac{2}{3N}$$

Relative variance of kinetic energy

$$\langle p^2 \rangle = \int d\mathbf{p} \cdot p^2 W(p) = \frac{3m}{b}$$

Second moment of $W(p)$

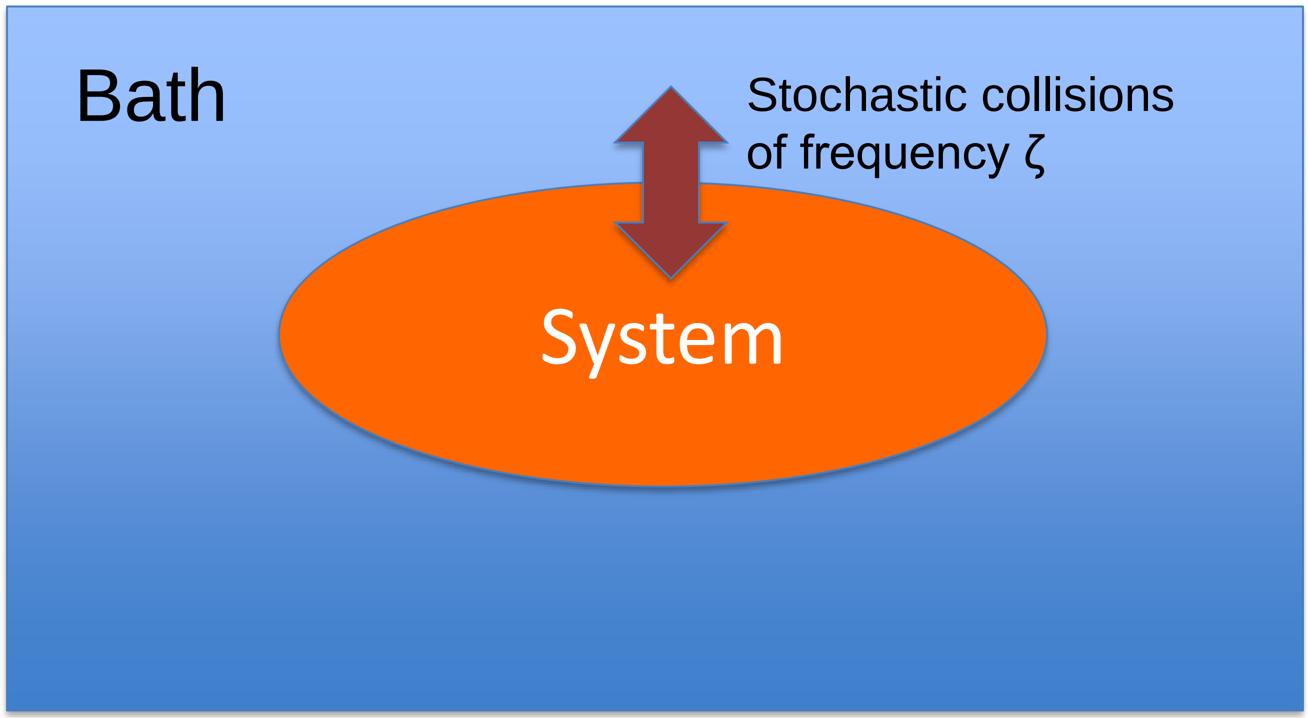
$$\langle p^4 \rangle = \int d\mathbf{p} \cdot p^4 W(p) = 15\left(\frac{m}{b}\right)^2$$

Fourth moment of $W(p)$

$$p^2 = \sum_a p_a^2$$

Instantaneous T_K fluctuates!

Andersen Thermostat



- Introduce a heat bath, and a coupling of the system with the bath.

- At specific time intervals (frequency = ζ), the system is subject to stochastic forces, which affect the energy of a system.

- In time, the system will evolve from constant energy state to constant energy states.

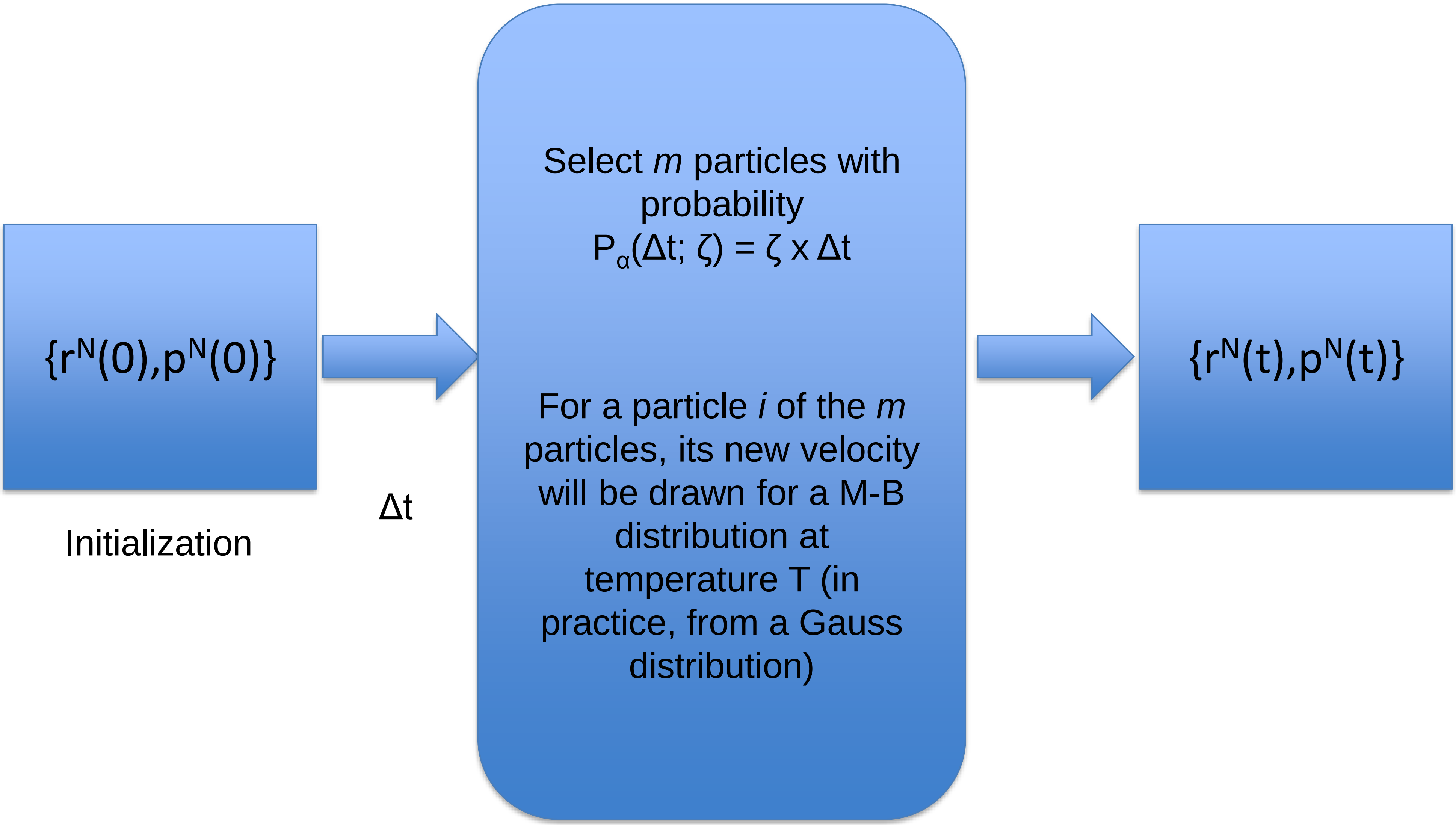
- Energy changes of the particle momenta are controlled by a Maxwell-Boltzmann (M-B) distribution.

- This ensures that all Boltzmann-accessible energies are visited and properly weighted.

- Technically, the system is propogated according to Newton's dynamic equations only between single collision events, while the momenta reshuffling implements a Monte Carlo move on the system. This corresponds to a Markov chain.

- Given enough time, the system evolves into an equilibirum distribution of momenta (M-B profile).

MD with Andersen Thermostat



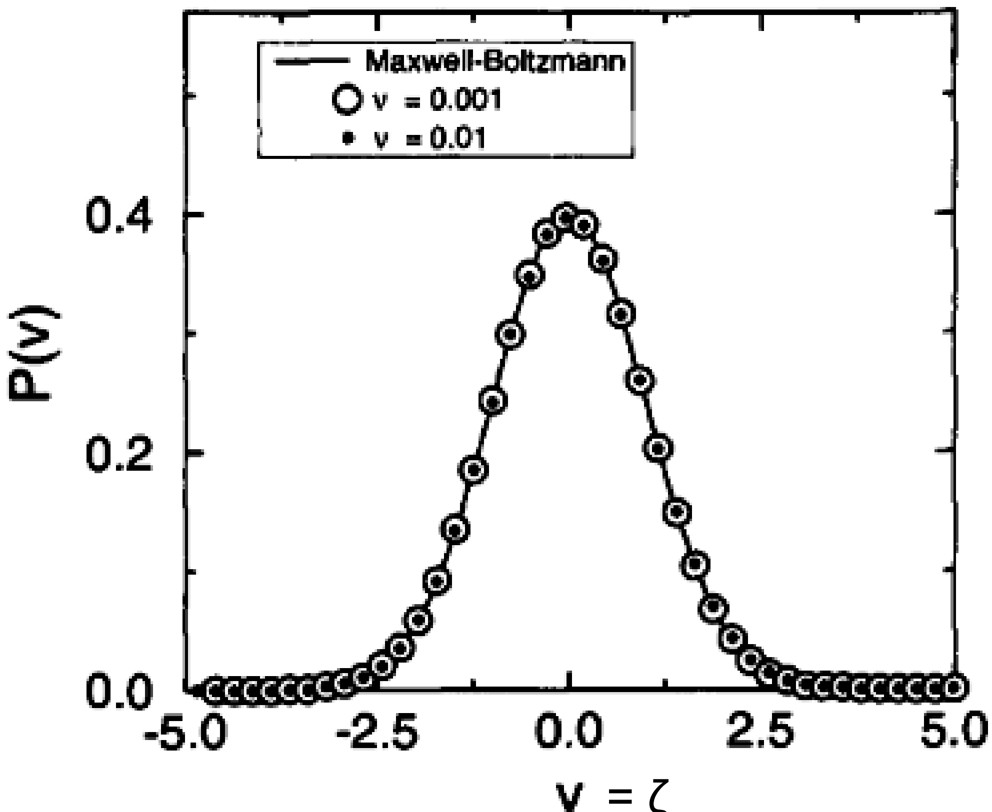
Possible Implementation

```
sigma = sqrt(temp)

do i=1, N
        if(ranf() < ζ x Δt)
                v(i) = gauss(sigma)
        endif
enddo
```

ranf(): [0,1] - random number generator
gauss(*sigma*): value from Gaussian distribution
of std. deviation *sigma*

Probability distribution profile

W(v) =

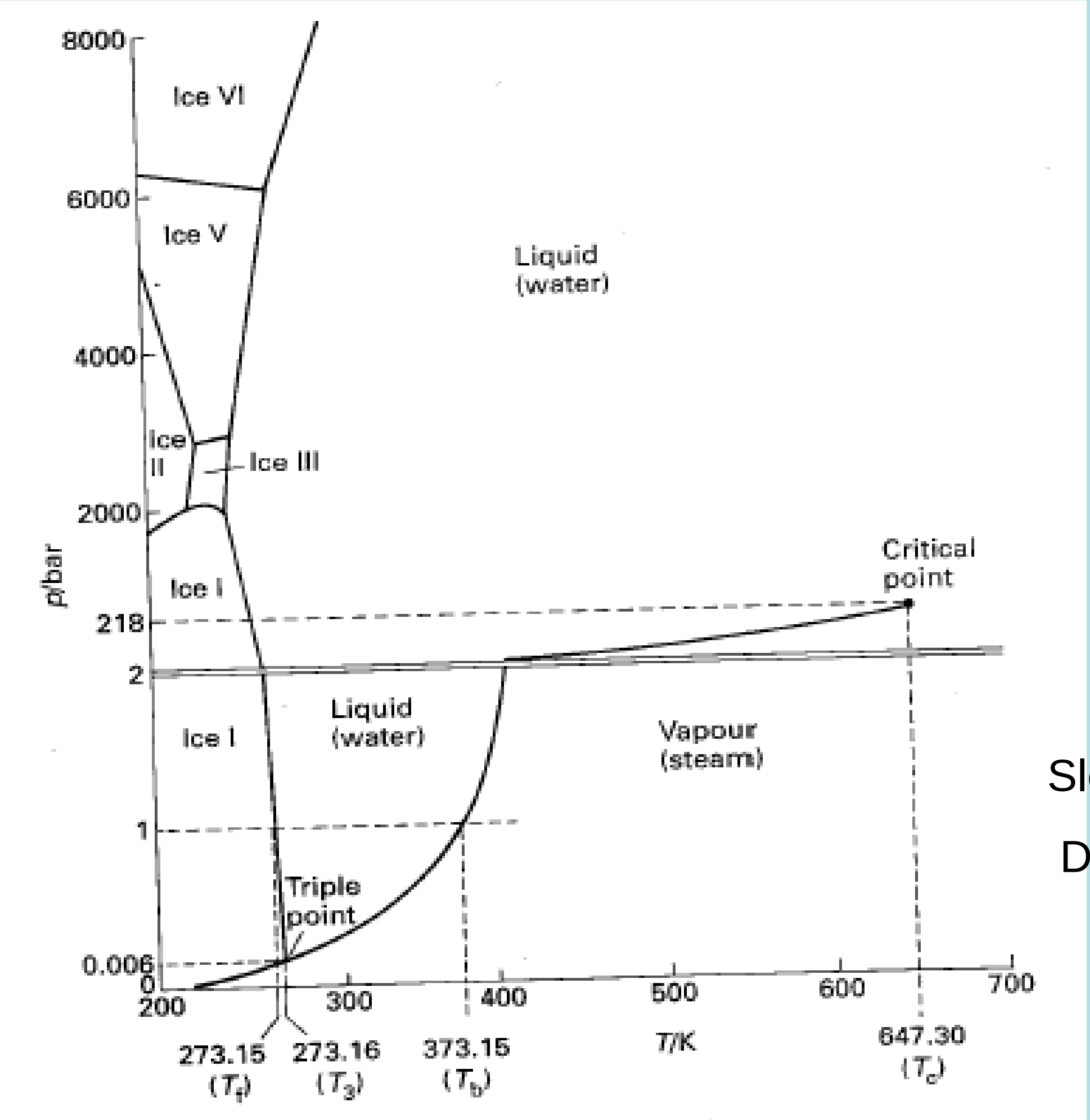


= ζ

Case Studies – Water

a) Liquid-solid (crystallisation);
b) solid-liquid & proton mobility at high pressure.

Water



Slope at triple point positive

Density of Ice < density of water.

Simulation of Water Crystallization by MD

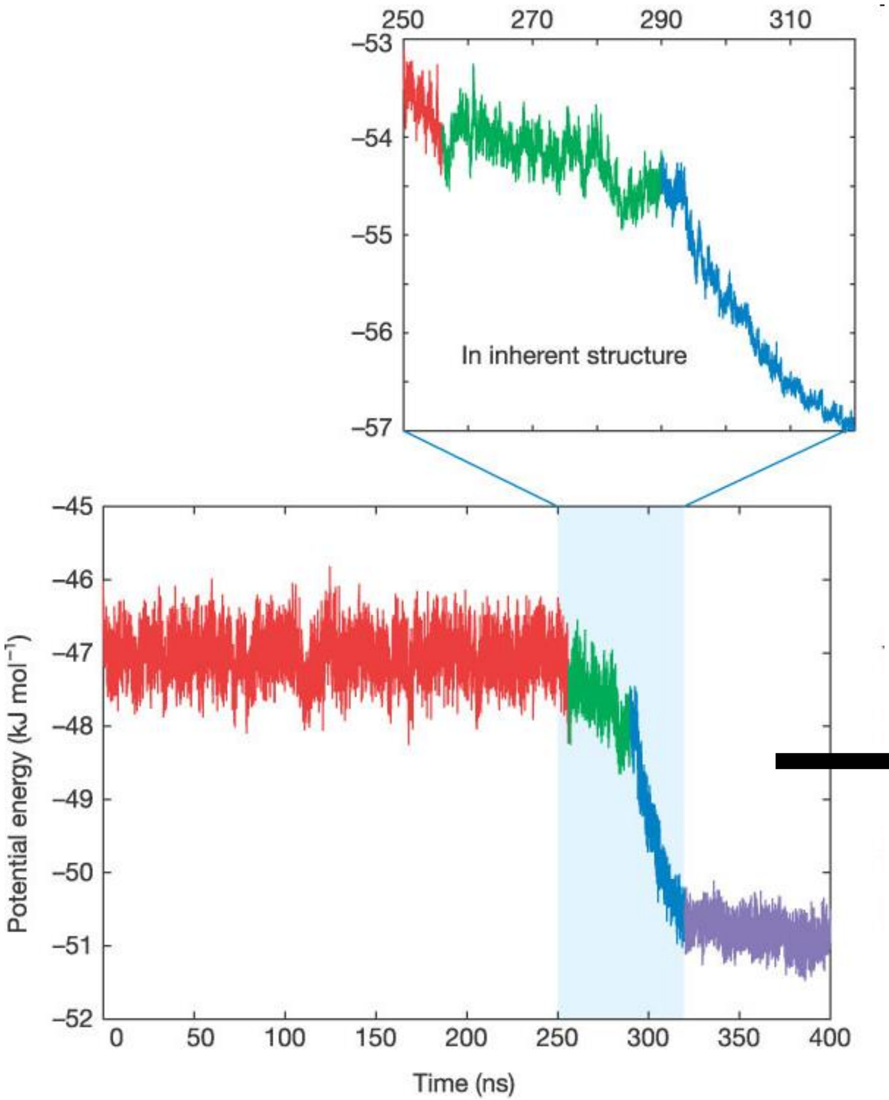
- Thermalization at high Temp
- Quenching to lower T (230 K)
- Supercooled state
- Time evolution, const-T, const-p
- 512 molecules
- Observation
- Order Parameter

Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing

Masakazu Matsumoto, Shinji Saito & Iwao Ohmine

Chemistry Department, Nagoya University, Chikusa-ku, Nagoya, Japan 464-8602

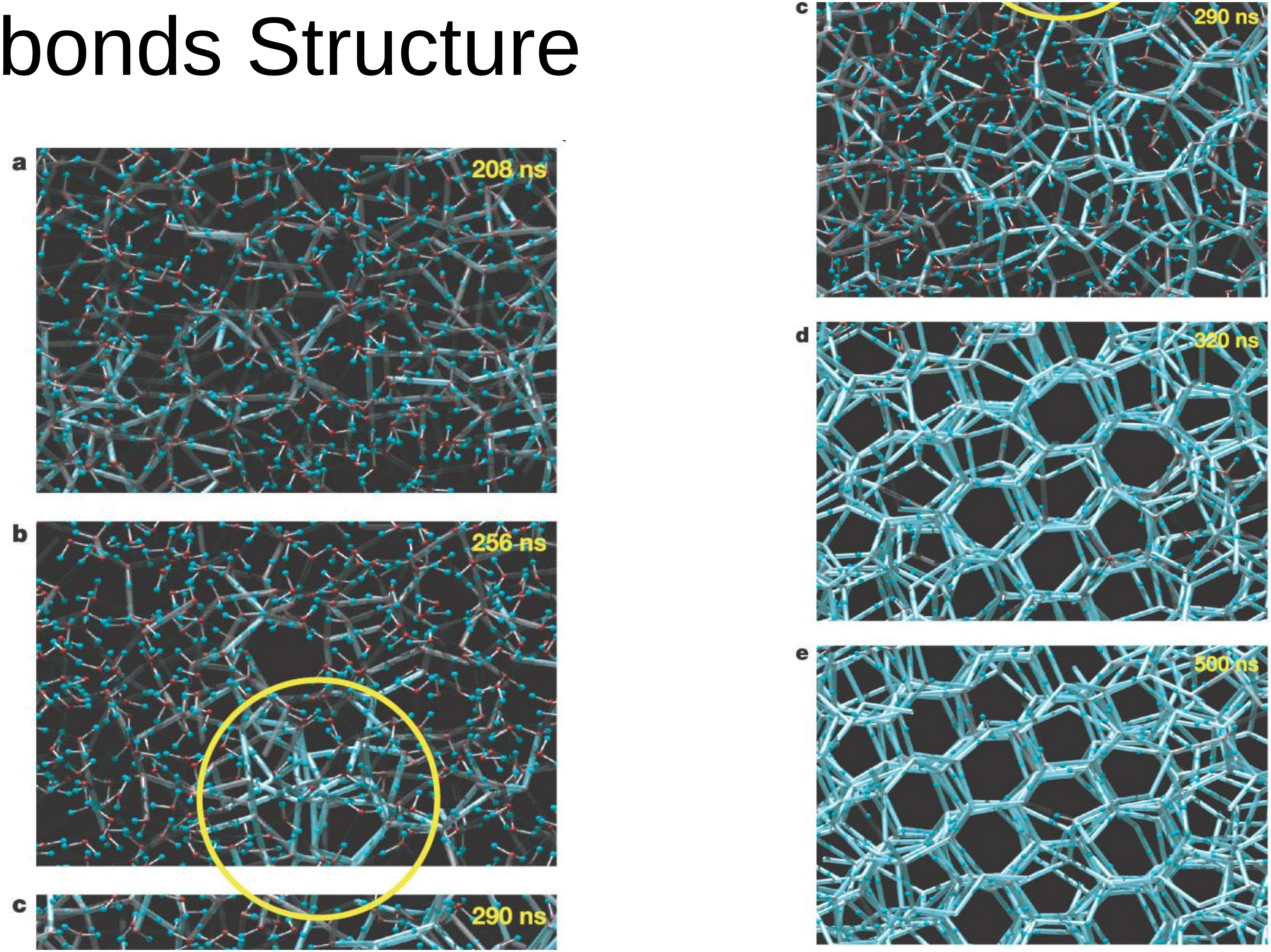
Inherent Structure



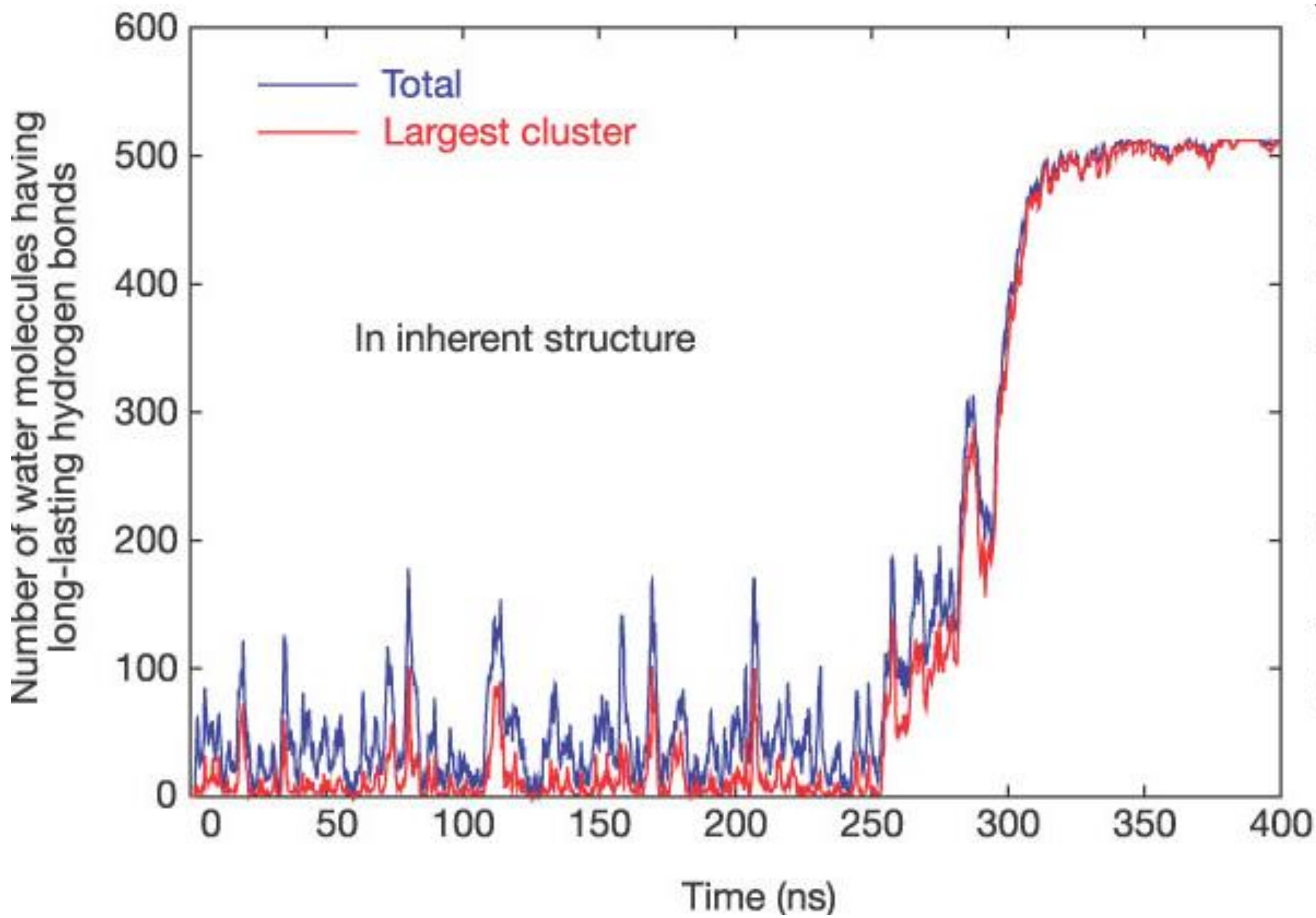
Samples of structures at 10 ps intervals,

Evolution of potential energy

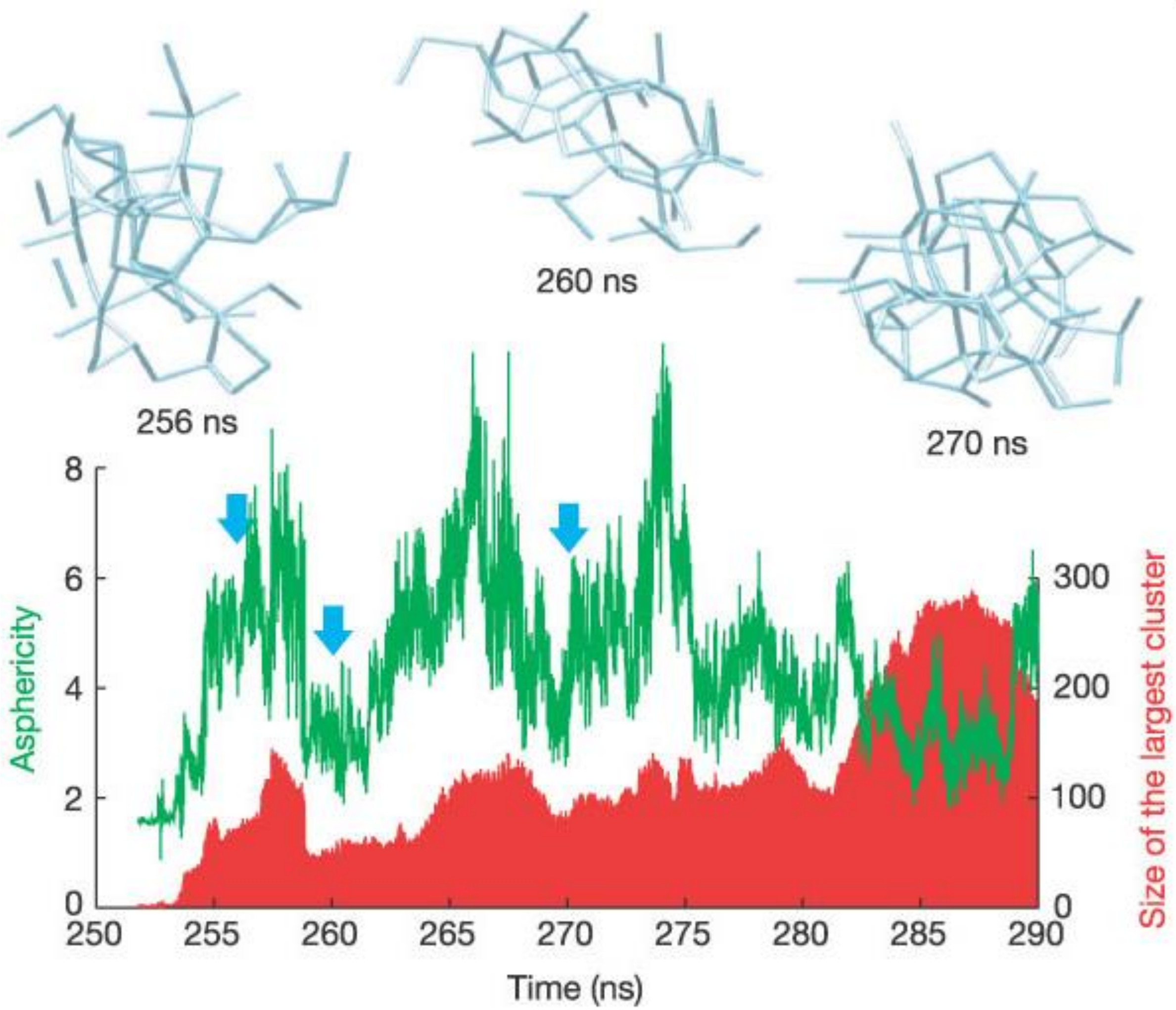
H-bonds Structure

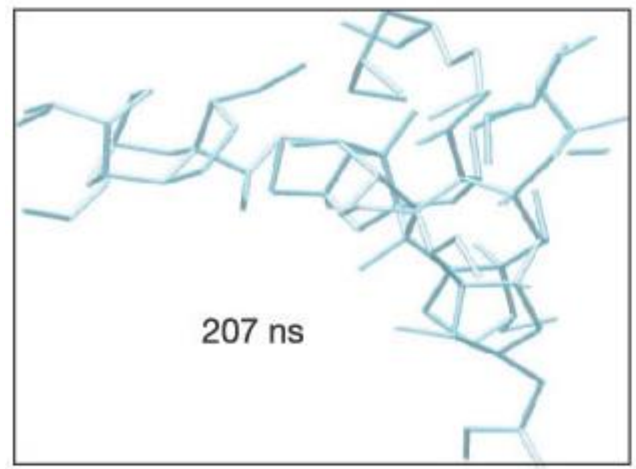


Monitoring of PT progress

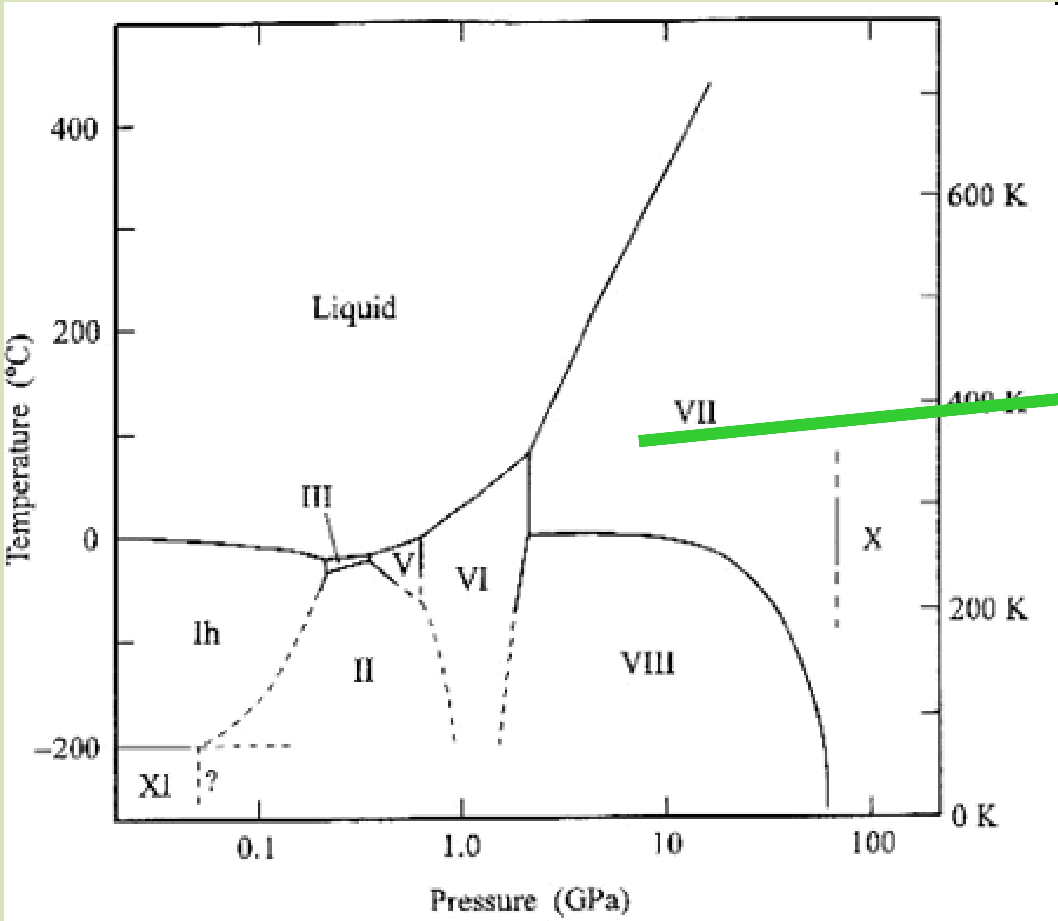


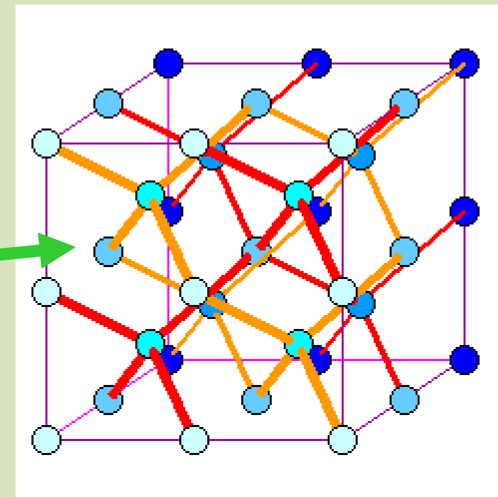
Local Structures





High pressure ice(s)





Melting of ice under pressure

Eric Schwegler*†, Manu Sharma‡, François Gygi*§, and Giulia Galli*‡

*Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550; and Departments of ‡Chemistry and §Applied Science, University of California, Davis, CA 95616

Communicated by Berni J. Alder, Lawrence Livermore National Laboratory, Livermore, CA, August 18, 2008 (received for review February 8, 2008)

The melting of ice under pressure is investigated with a series of first-principles molecular dynamics simulations. In particular, a two-phase approach is used to determine the melting temperature of the ice-VII phase in the range of 10–50 GPa. Our computed melting temperatures are consistent with existing diamond anvil cell experiments. We find that for pressures between 10 and 40 GPa, ice melts as a molecular solid. For pressures above ≈45 Gpa, there is a sharp increase in the slope of the melting curve because of the presence of molecular dissociation and proton diffusion in the solid before melting. The onset of significant proton diffusion in ice-VII as a function of increasing temperature is found to be gradual and bears many similarities to that of a type-II superionic solid.

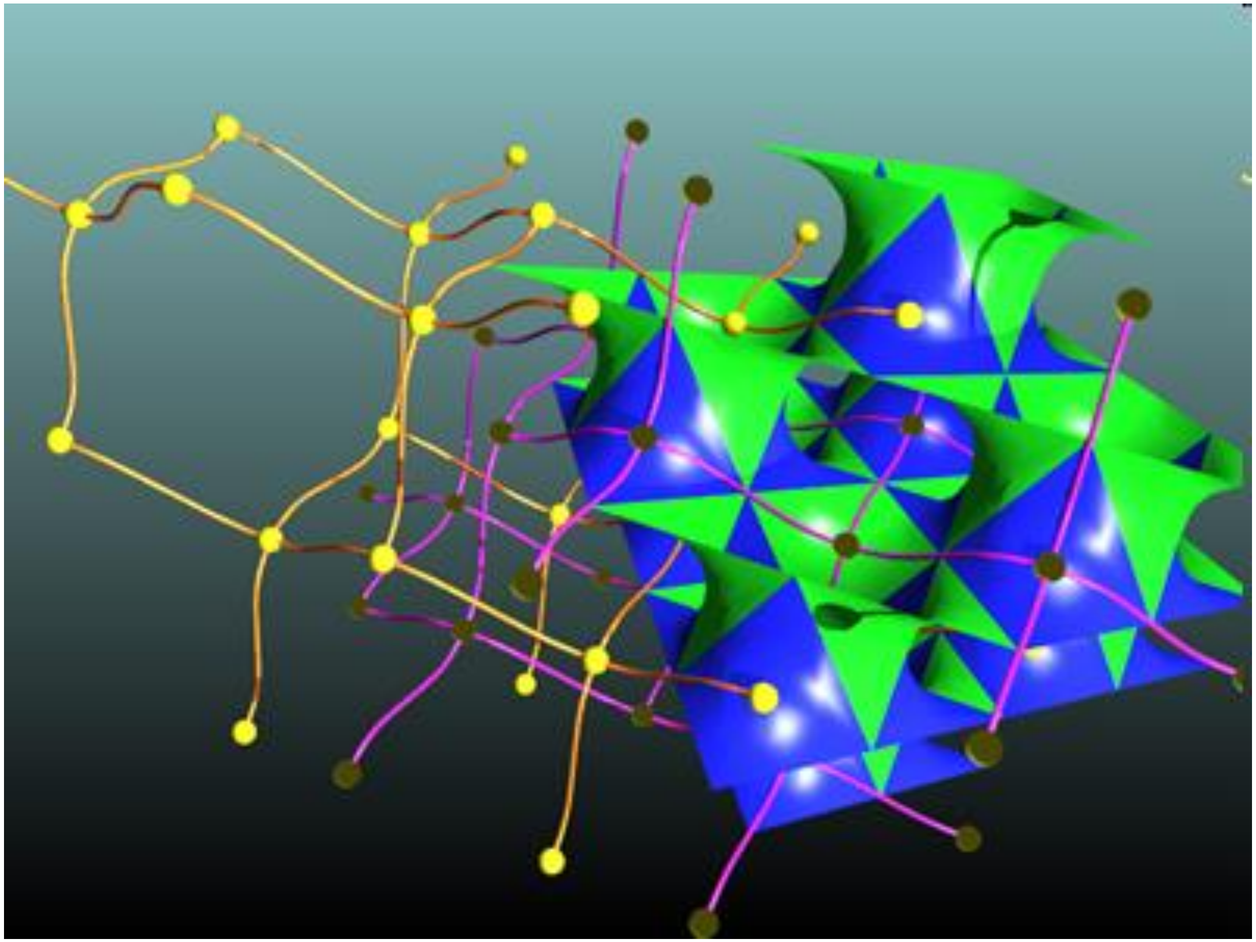
Ice: solid-liquid (VII)

- Determination of the melting line at high pressures;
- Straighforward MD possibly problematic, due to superheating or undercooling effects;
- Large discrepancies between theory/experiment;
- Melting process as a molecular solid, or dissociation ?
- Which level of theory is necessary here ?
- Range: 10-50 GPa

Preparing the model

- Ice VII is described as 2 interpenetrating cubic (diamond-like nets).
- Oxygen are places on a bcc lattice, hydrogen are disordered.
- Prepare the O on a *bcc* lattice, distribute H in a random manner, however:
  - Use Monte Carlo moves to minimize the dipole, obey Pauling ice rules (two-in, two-out);

Interpenetrating networks

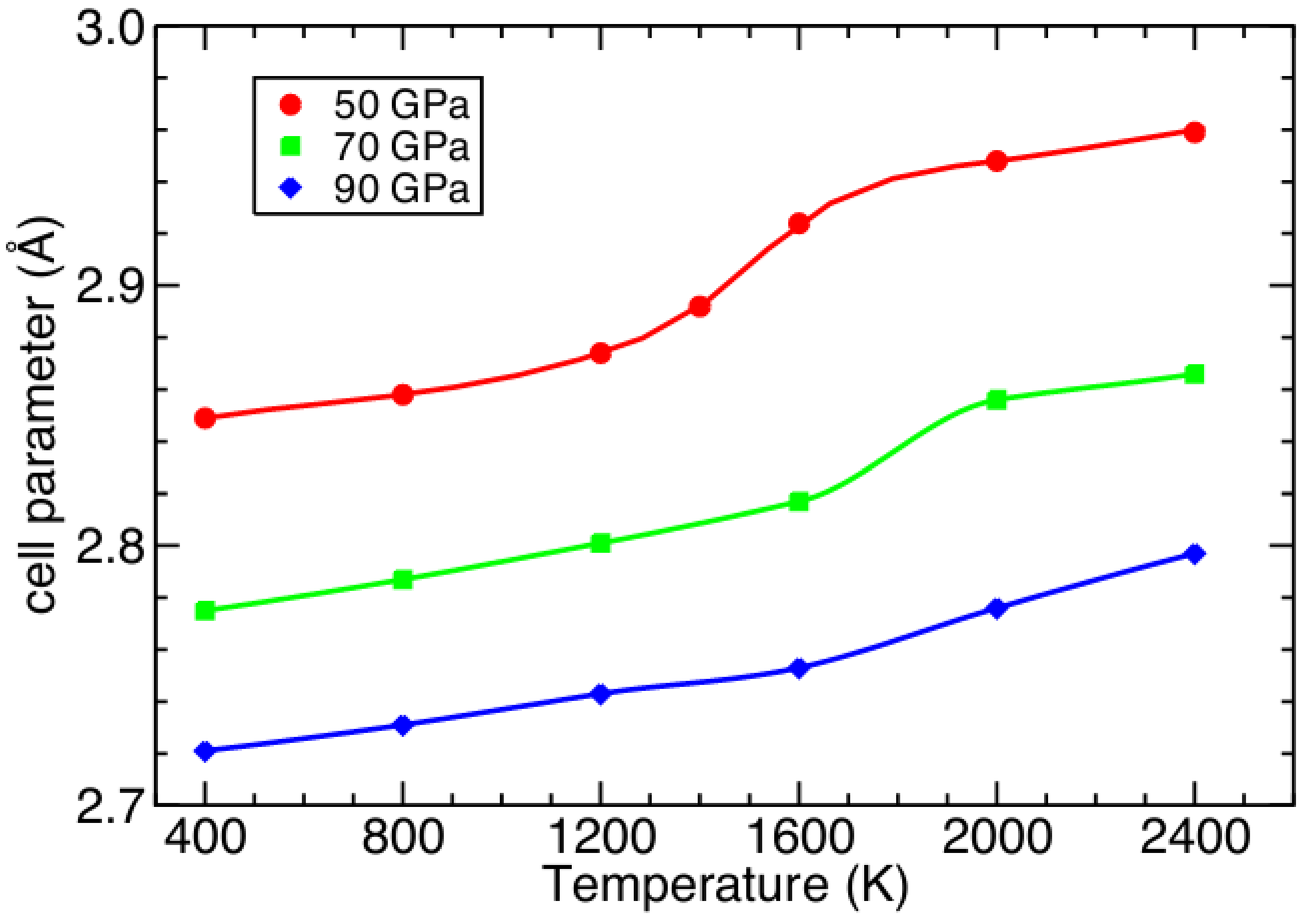


http://epinet.anu.edu.au/infinite_tiles/s2224_tree

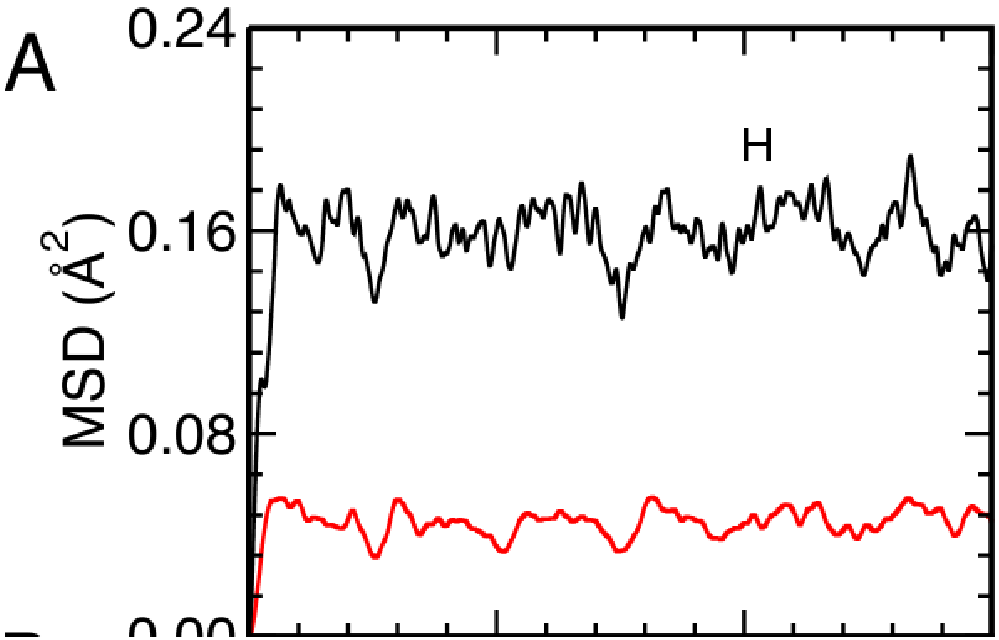
Describing Ice VII & water

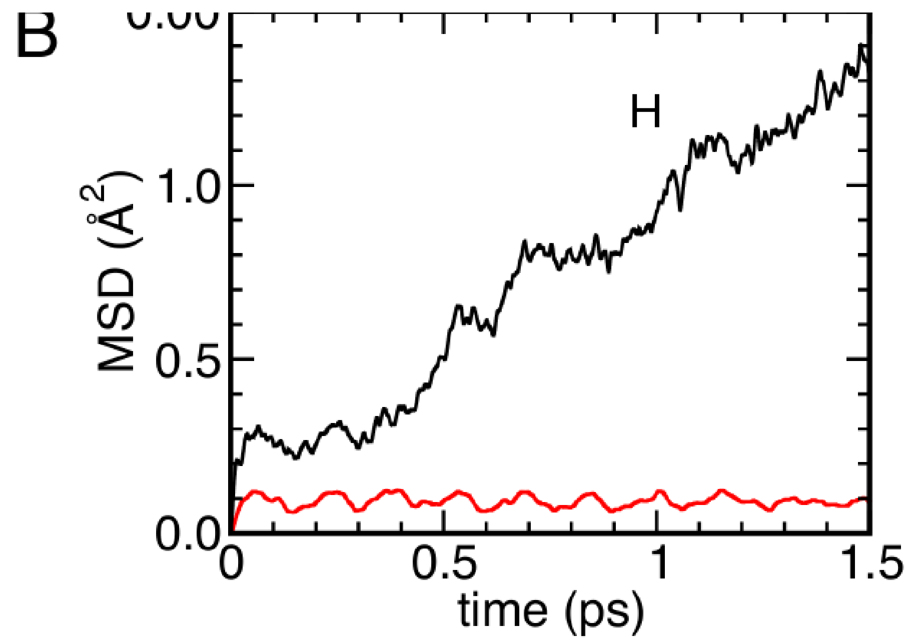
- Start simulation from the „pure“ phases. Prepare corresponding boxes, propagate in time at different temperatures and pressure, in cold-hot, or hot-cold runs;
- Trace some properties that may change as the effect of pressure/temperature;
- A discontinuity is often the signature of a structural change.

Lattice parameters of ice VII under p



Different O/H behavior





1,200 K

1,400 K

Simulation of melting/solidification

- Start from configurations obtained by simulating at identical conditions water and ice;
- Prepare a simulation box with half ice and half water → phase coexistence;
- Two phase simulation „method“;
- Equilibrate, to achieve an interfacial structure;
- Propagate at different temperatures;
- Successively bisect the melting temperature

Phase diagram

$$\mu^{(\alpha)}(p,T) = \mu^{(\beta)}(p,T)$$
$$\mu^{(\alpha)}(p,T) = \mu^{(\gamma)}(p,T)$$
$$\mu^{(\beta)}(p,T) = \mu^{(\gamma)}(p,T)$$

PT -> intersection of two planes

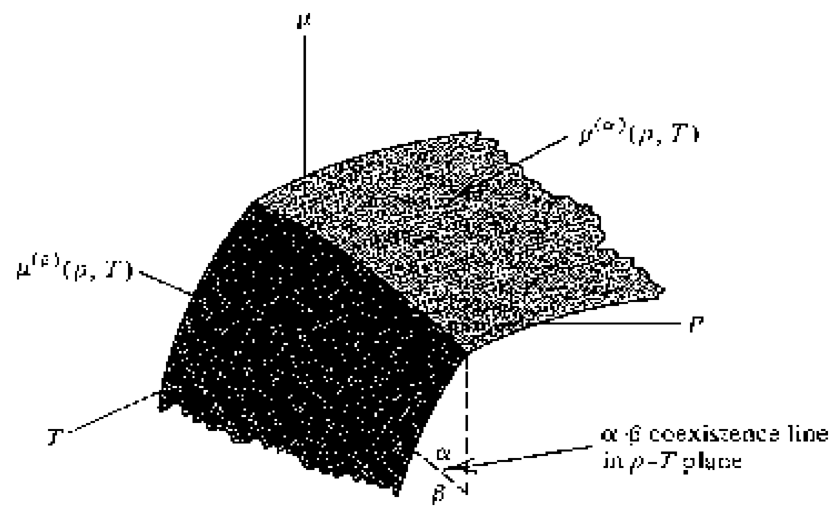


Fig. 2.5. Chemical potential surfaces for two phases.

Phase diagram

$$\mu^{(\alpha)}(p,T) = \mu^{(\beta)}(p,T)$$

$$d\mu = -sdT + vdp$$

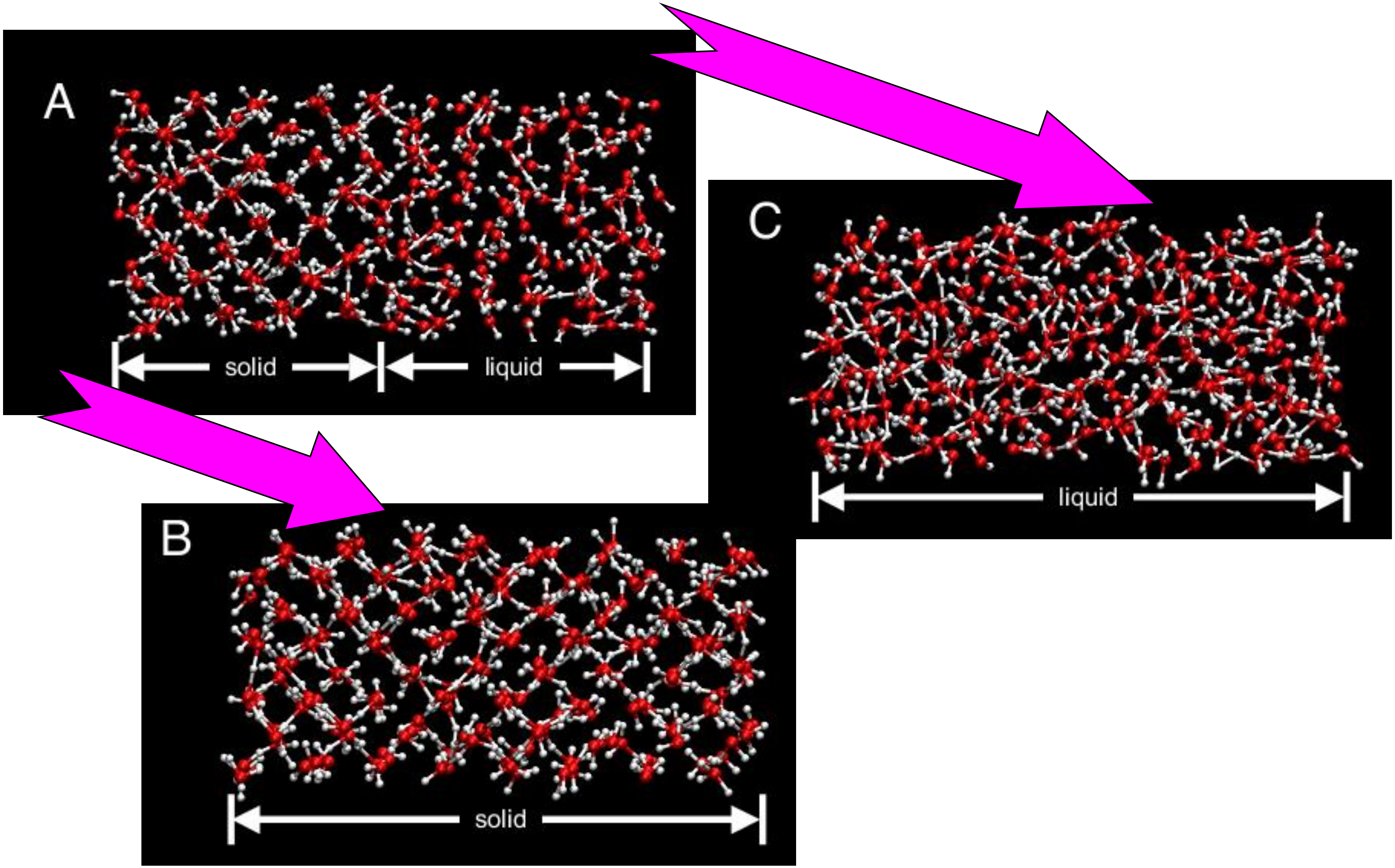
$$-s^{\alpha}dT + v^{\alpha}dp = -s^{\beta}dT + v^{\beta}dp$$

$$\frac{dp}{dT} = \frac{\Delta s(T)}{\Delta v(T)}$$

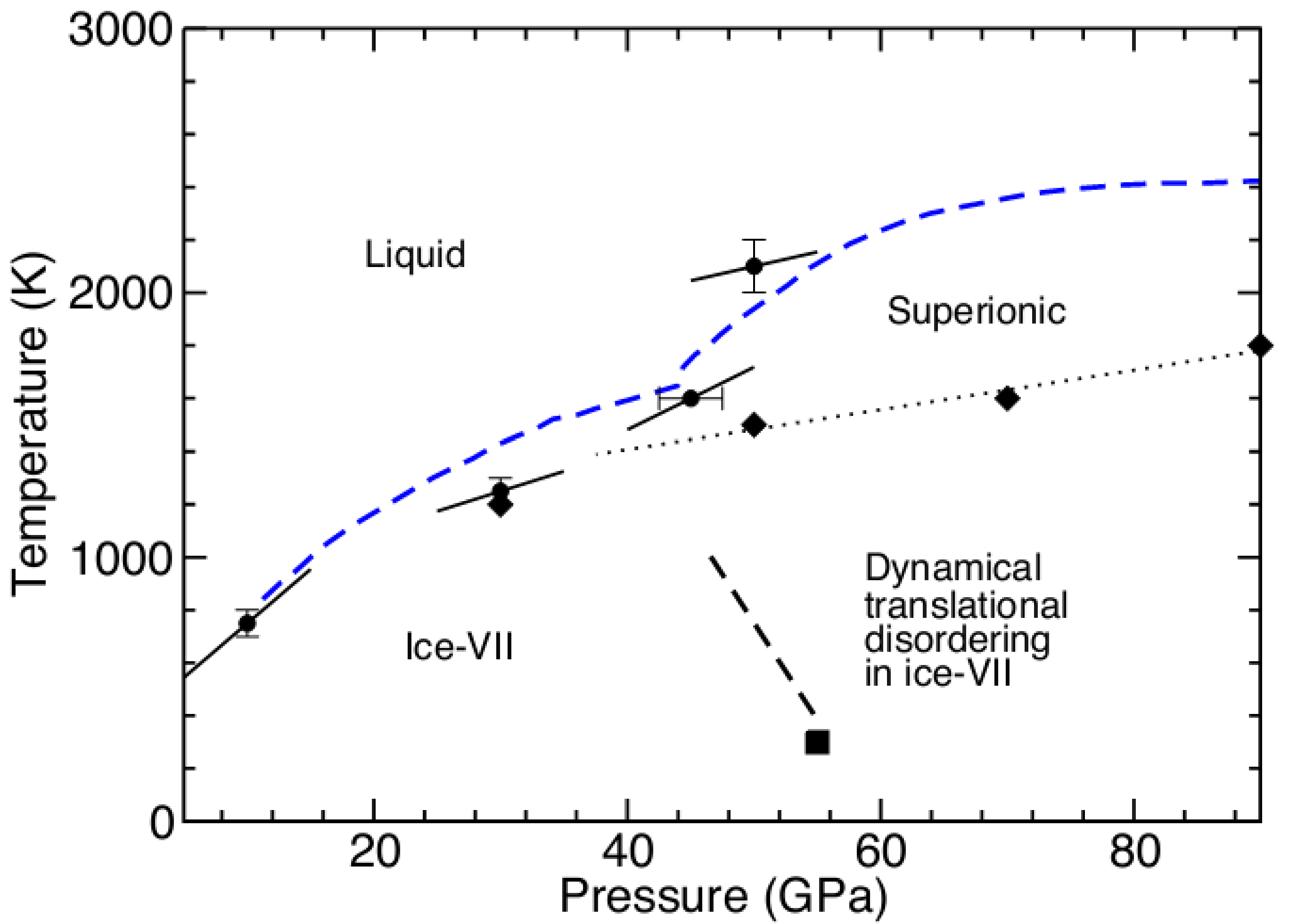
Clausius-Clapeyron Equation

In general, for solid-liquid PT $S_l > S_s$, $V_l > V_s$,

Two-phases simulations



Phase Diagram



Explicit MD

WORKSHOP 1

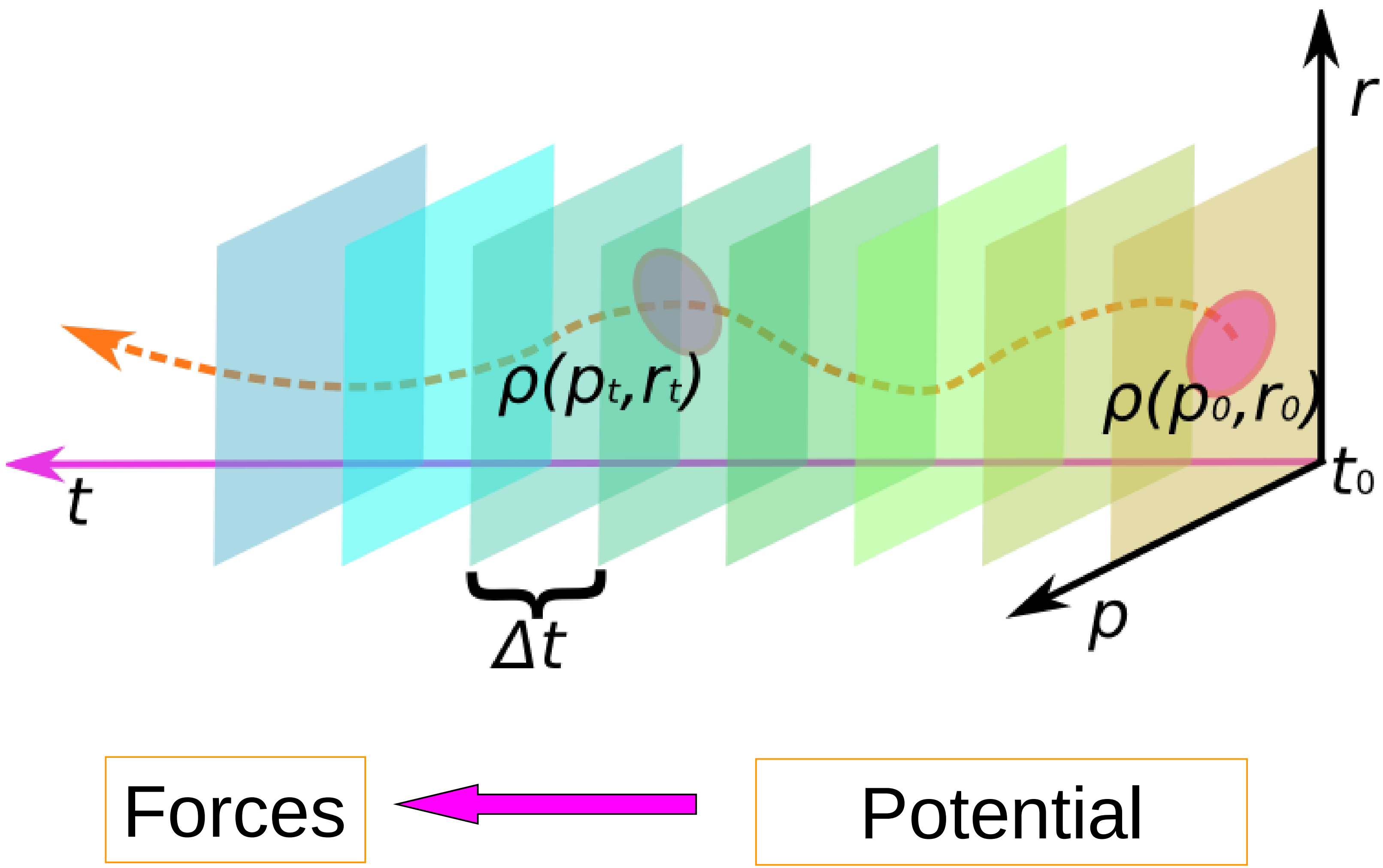
SIESTA

- Linear-scaling DFT based on NAOs (Numerical Atomic Orbitals)
- Main Reference*: P. Ordejon, E. Artacho & J. M. Soler , Phys. Rev. B 53, R10441 (1996) J. M.Soler et al, J. Phys.: Condens. Matter* ***14****, 2745 (2002)*
- *Spanish Initiative for Electronic Structure Calculations with Thousands of Atoms*

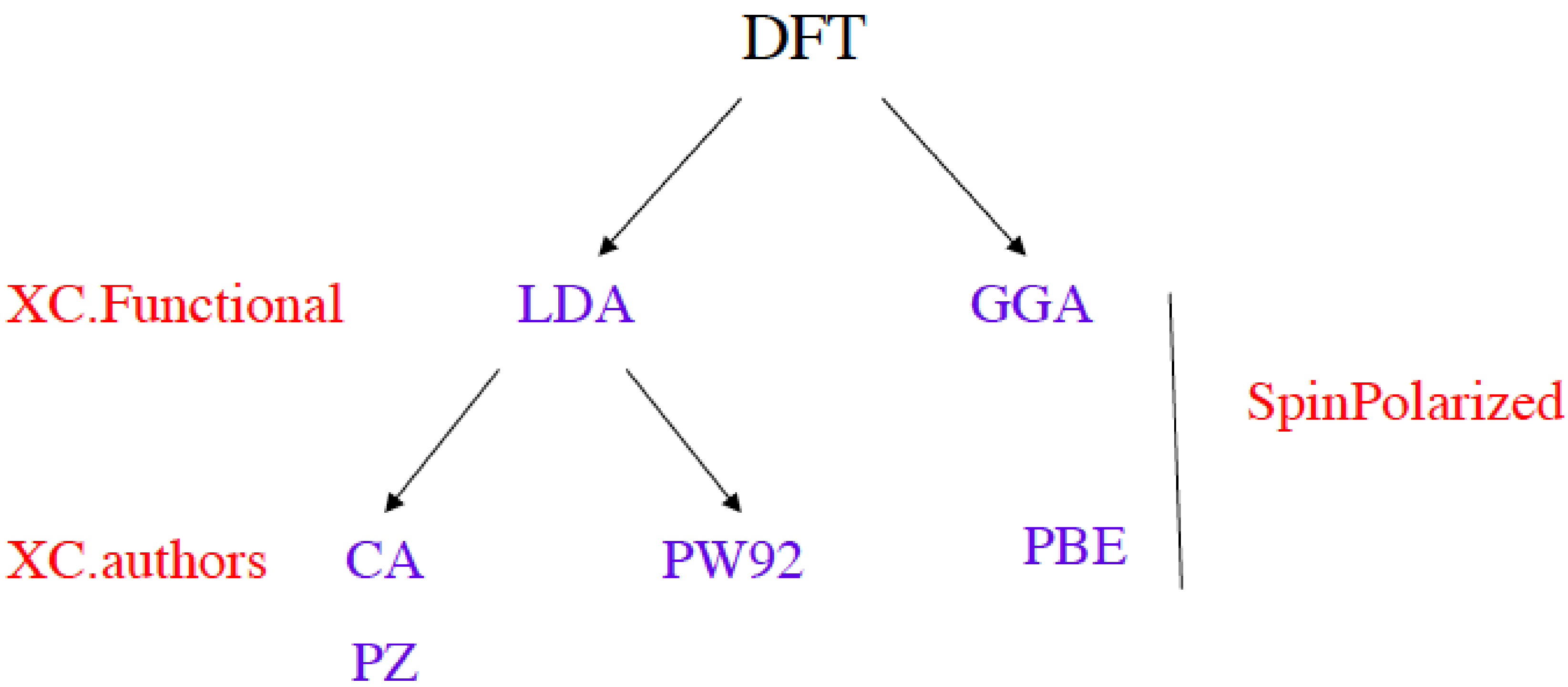
Scope

- BO (forces)
  - Molecular Dynamics & Optimisation
- DFT – LDA/GGA (also LDA/GGA +U)
- Pseudopotentials
  - No explicit treatment of core electrons – faster, however tricky…
- Numerical orbitals & numerical evaluation of matrix elements

MD, Trajectory



Functional



DFT ≡ Density Functional Theory

LDA ≡ Local Density Approximation

GGA ≡ Generalized Gradient Approximation

VDW ≡ van der Waals

CA ≡ Ceperley-Alder

PZ ≡ Perdew-Zunger

PW92 ≡ Perdew-Wang-92

PBE ≡ Perdew-Burke-Ernzerhof

DRSLL ≡ Dion et al (VdW)

Specialised functionals

VOLUME 92, NUMBER 24 PHYSICAL REVIEW LETTERS week ending 18 JUNE 2004

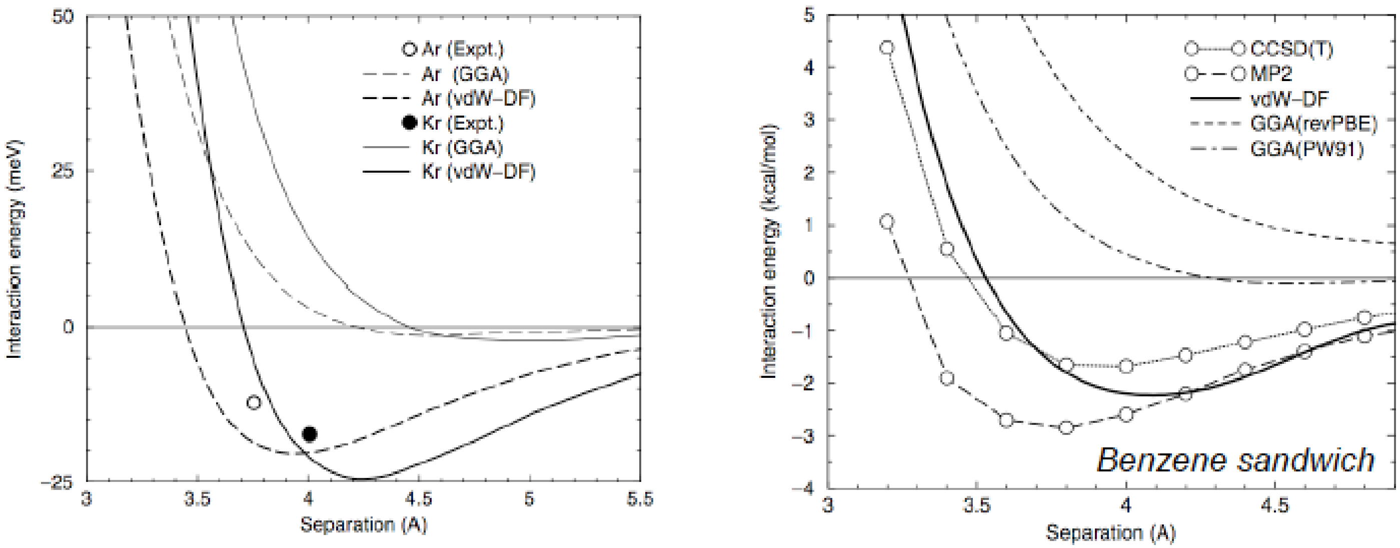
Van der Waals Density Functional for General Geometries

M. Dion,[1] H. Rydberg,[2] E. Schröder,[2] D. C. Langreth,[1] and B. I. Lundqvist[2]

[1]*Center for Materials Theory, Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854-8019, USA*

[2]*Department of Applied Physics, Chalmers University of Technology and Göteborg University, SE-412 96 Göteborg, Sweden*

(Received 30 January 2004; published 16 June 2004)



Practical strategy, not perfect, successful for some system, would fail in others.
Rather a local correction than a non-local functional.

Solution method

0: Start from the atomic coordinates and the unit cell

$$\{\vec{R}\}_N \; \{\vec{a}\}$$

1: Compute H, S (Order N):

Hamiltonian (H), Overlap (S) matrices

$$(H - \varepsilon S)C = 0$$

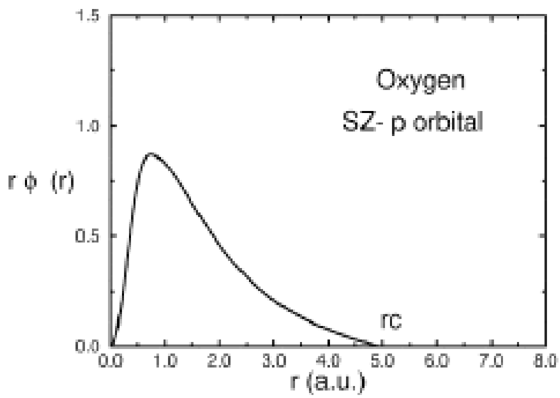
2: SolutionMethod

diagon

Order-N

Execution time

Hard confining potentials



Fireballs

O. F. Sankey & D. J. Niklewski,
Phys. Rev. B 40, 3979 (1989)

BUT:

A different cut-off radius for each orbital

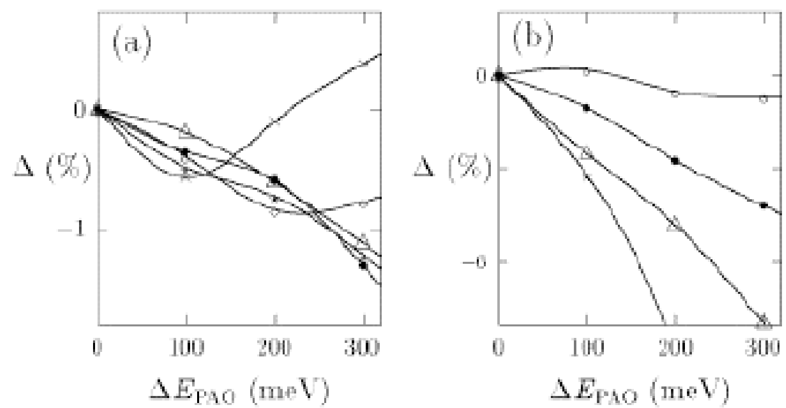
A single parameter

Energy shift

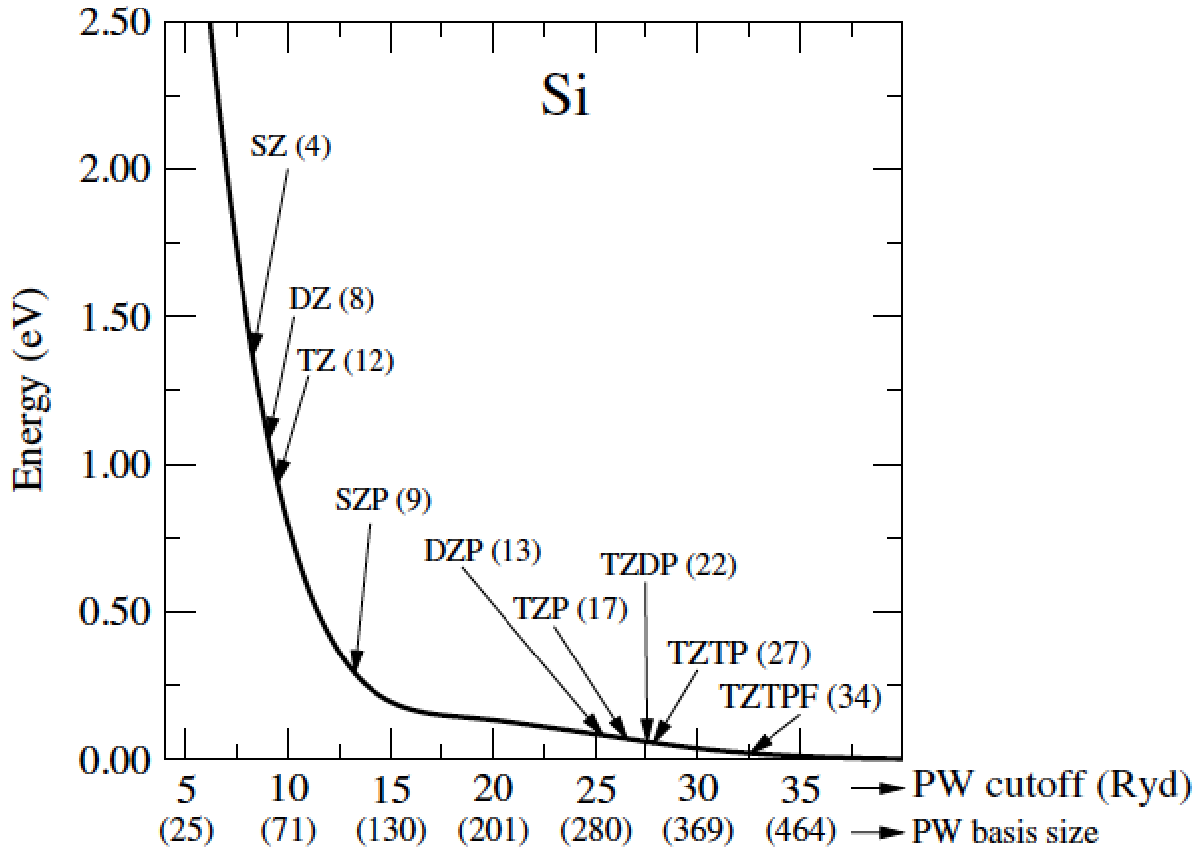
E. Artacho et al. *Phys. Stat. Solidi (b) 215, 809 (1999)*

Convergence vs Energy shift of

Bond lengths *Bond energies*

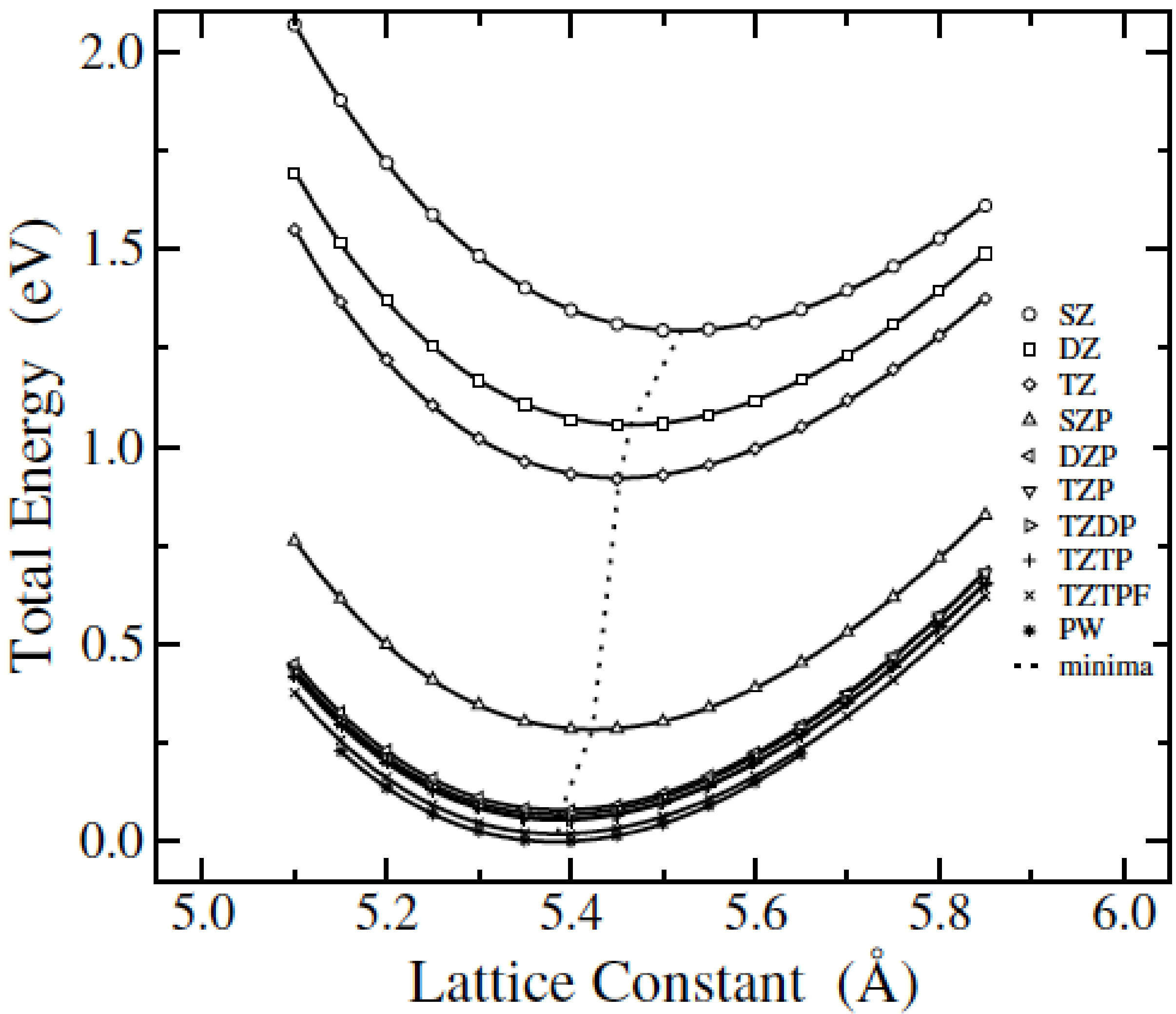


Basis - size



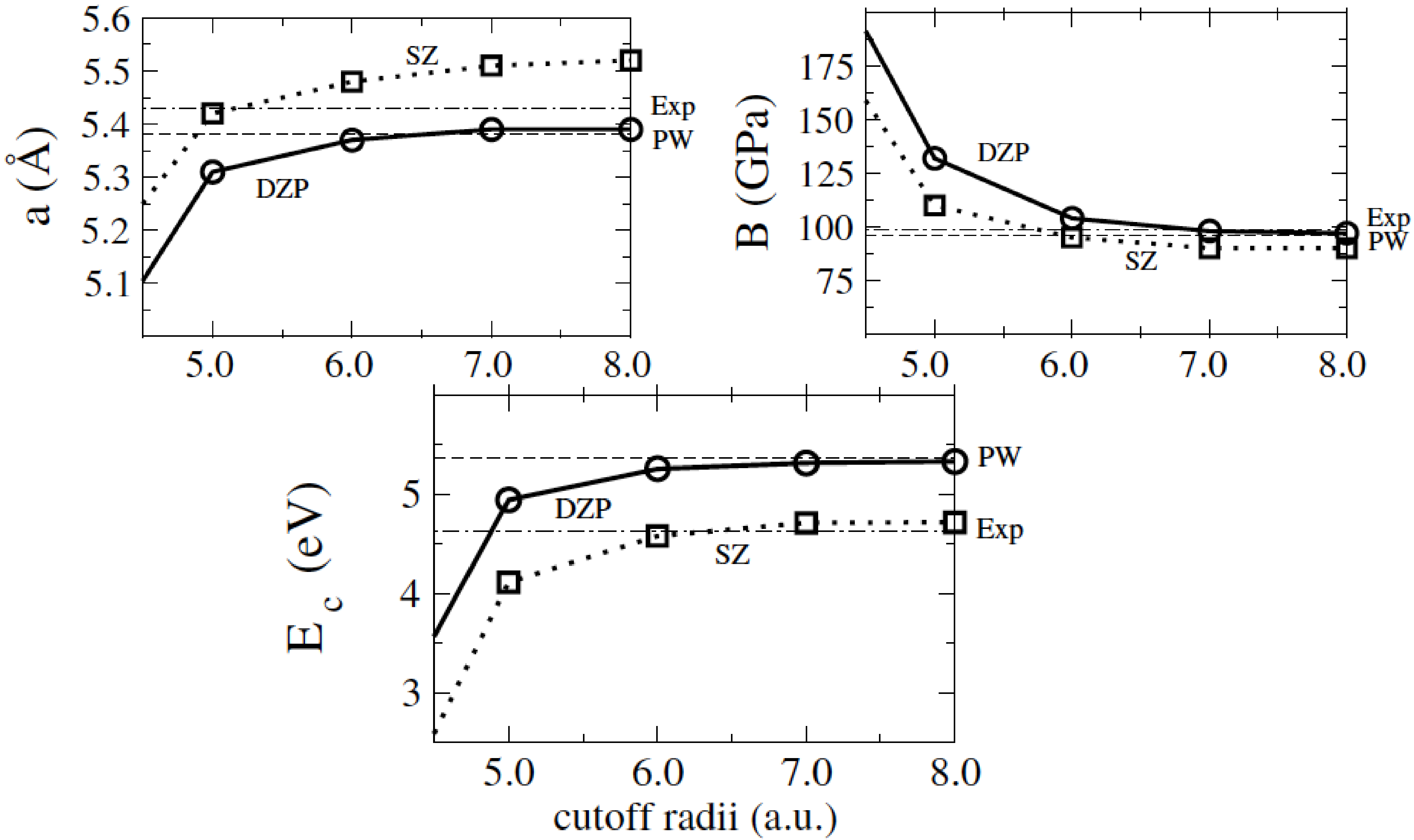
Comparison LCAO with PW –
from J M Soler, 2002 J. Phys.: Condens. Matter 14 2745

Location of the minimum



Si

Properties vs cutoff radius



Different basis size

Basis	a (Å)	B (GPa)	E_c (eV)
SZ	5.521	88.7	4.722
DZ	5.465	96.0	4.841
TZ	5.453	98.4	4.908
SZP	5.424	97.8	5.227
DZP	5.389	96.6	5.329
TZP	5.387	97.5	5.335
TZDP	5.389	96.0	5.340
TZTP	5.387	96.0	5.342
TZTPF	5.385	95.4	5.359
PW	5.384	95.9	5.369
LAPW	5.41	96	5.28
Expt	5.43	98.8	4.63

Resolution

- Start with SZ
- Lower values of the numerical mesh (150-200 Ry)
- No k points if possible (gamma point calculations for large structures, critical if cell parameters are small)
- Fermi smearing (for metals, but may also help convergence)

What should be done (in principle)

- Prepare pseudopotential(s)
- Test pseudopotentials
- Optimize the basis
- Critical choice of XC functional(s)
- Calculate SZ, DZ, DZP (TZP ??)
- Compare with a full potential method (or PW method)
- Pedantic convergence test (mesh cutoff, k points, basis size)

What it is done (in practice)

- There are databases of pseudopotentials, and standard choices of the basis (also within SIESTA).
- SZ gives a good initial guess, it is also fast
- There are also good mesh cutoff values
- Normally some experience with a particular element helps
- A routine user would accumulate a “personal” database of pseudos & basis…

Amoeba

Nonlinear Optimisation

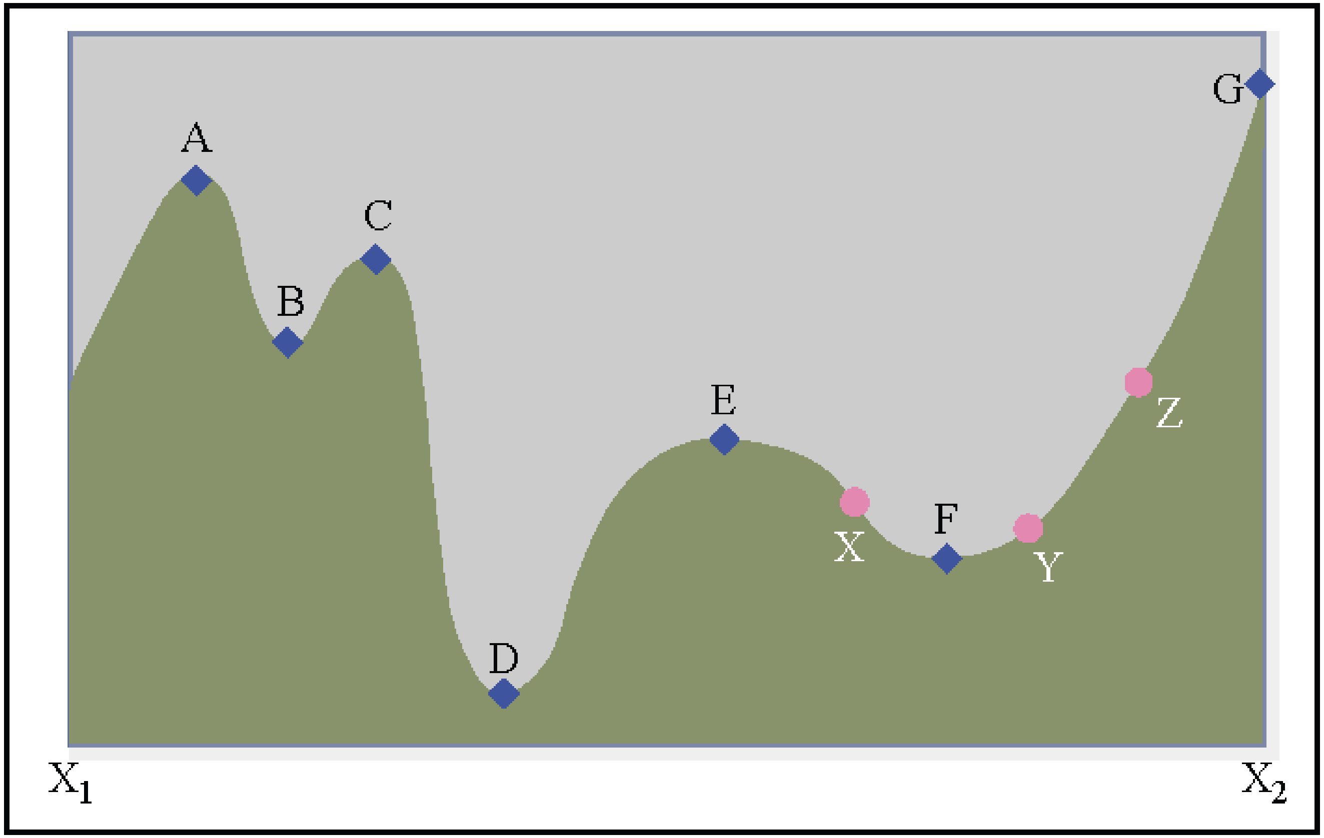


Figure by MIT OpenCourseWare.

https://ocw.mit.edu/courses/civil-and-environmental-engineering/1-204-computer-algorithms-in-systems-engineering-spring-2010/lecture-notes/MIT1_204S10_lec22.pdf

non-linear algorithms

- Amoeba
  - Nelder-Mead method
  - Direct solution of non-linear optimisations problems
  - No derivatives
  - Variable/adaptable step size based on function value
  - Simplest, most robust, slowest
  - No assumprions about function needed, “universal”
- CG
  - Requires function, first derivative (gradient)
  - Step size adapts as algorithm advances

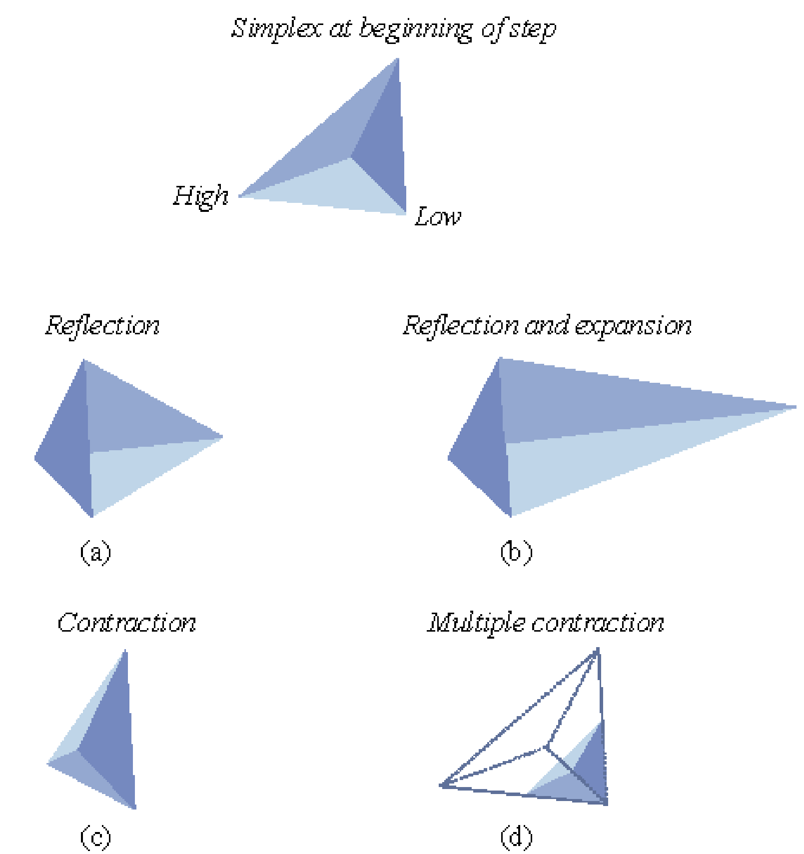
Amoeba algorithm

- Requires function evaluation only
- Less efficient compared to derivative-based, but more robust
- Short, compact implementation
- Appropriate if derivative are difficult (or possibly inexact)
- “Crawls downhill with no assumptions about function”

Amoeba steps

- Definition of a n+1 dimensional simplex for n-dimensional problem (a tetrahedron in 3D – 4 vertices).
- Every vertex is a function value.
- Starting points can be guessed values or radom choices (random choice of parameters from which function value is calculated)
- Minimisation steps have a geometric mapping

Amoeba Steps



Possible outcomes for a step in the downhill simplex method

Figure by MIT OpenCourseWare.

Simplex: geometric changes

- Move points where function is highest (reflection)
- Where/if function flat, expand, then reflect (change/adapt step)
- Outcome of each step:
  - Contraction in some directions
  - Overall contraction
- Termiantion: values at vertices within given threshold (“zero” volume)